

THE PROBLEM OF MORAL AMBIVALENCE

REVISITING HENRY SIDGWICK'S THEORY OF RATIONAL BENEVOLENCE AS A BASIS FOR MORAL REASONING, WITH REFERENCE TO PRENATAL ETHICAL DILEMMAS

Submitted by Rachel Helen Addison to the University of Exeter as a thesis for the degree of
Doctor of Philosophy in Theology, in July 2016

This thesis is available for library use on the understanding that it is copyright material, and
that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified, and that
no material has previously been submitted and approved for the award of a degree by this or
any other University.

Signature:.....

THE PROBLEM OF MORAL AMBIVALENCE

REVISITING HENRY SIDGWICK'S THEORY OF RATIONAL BENEVOLENCE AS A BASIS FOR MORAL REASONING, WITH REFERENCE TO PRENATAL ETHICAL DILEMMAS

RACHEL ADDISON

ABSTRACT

This thesis addresses the conflict traditionally found within moral philosophy between deontological and utilitarian schools of thought. Using the example of the serious moral ambivalence experienced by individuals who are deciding whether to end or continue a difficult pregnancy, it is argued that this ambivalence is the result of both absolute principles (such as the intrinsic value of human life) and outcome based considerations (such as the desire to avoid causing pain and suffering) appearing to be morally reasonable, while also being fundamentally opposed: Each course of action is at once morally defensible on the basis of its own reasonableness, and, conversely, reprehensible due to the reasonableness of the other. This lived experience of moral ambivalence is directly reflected by the tension between deontology and utilitarianism as it occurs at the moral philosophic level, where the deontological emphasis on the unconditional rightness of certain principles is seen to be at irreconcilable odds with the utilitarian emphasis on the attainment of certain ends. The thesis' central claim is that such ambivalence strongly indicates that human morality is neither exclusively one type or the other, and that both types of moral property are in fact reasonable, and thus have moral value. It is theorised that accounting for this dual reasonableness would lead to the most accurate and helpful representation of the human moral experience – but that the philosophic 'divide' between the two types of principle has led to an either/or situation, which has largely prevented this sort of understanding from being developed.

The thesis argues that Victorian philosopher Henry Sidgwick developed a view in which neither deontological nor utilitarian principles can be fully realised without reference to the other, precisely on the basis that both can be found to be ultimately rational. This thesis aims to revitalise that theory – represented by the term 'Rational Benevolence' - to show that Sidgwick reconciled the divide between absolute and end based principles in such a way that the relationship between them becomes a 'synthesis'. In this synthesis, deontological and utilitarian concepts are both seen as essential components of morality, that combine to form a dynamic whole in which the value of each principle is both indicated *and* naturally limited by the value of the other, on account of their respective rationalities. It is argued that this provides a more comprehensive understanding of the reality of the human moral experience, and better moral justification for *either* course of action in situations of complex and sensitive ethical decision making.

THE PROBLEM OF MORAL AMBIVALENCE

REVISITING HENRY SIDGWICK'S THEORY OF RATIONAL BENEVOLENCE AS A BASIS FOR MORAL REASONING, WITH REFERENCE TO PRENATAL ETHICAL DILEMMAS

RACHEL ADDISON

TABLE OF CONTENTS

TITLE PAGE/DECLARATION.....	1
ABSTRACT.....	2
TABLE OF CONTENTS.....	3
0. INTRODUCTION.....	9
0.1. The Problem of Moral Ambivalence: 'The Parental Predicament'...	9
0.2. The Significance of Moral Ambivalence: Two Types of Moral Consideration as 'Reasonable'	13
0.3. Moral Ambivalence, and the Problem of the 'Divide' in Moral Philosophy.....	16
<i>0.3.a. The 'Divide' in Moral Philosophy; Ambivalence Between Theories.....</i>	<i>17</i>
<i>0.3.b. The Problem of the 'Divide'; Autonomy and the Illusion of Choice.....</i>	<i>18</i>
0.4. Challenging the Divide: Reconciliation and Synthesis Between Deontological and Utilitarian Principles of Morality.....	20
<i>0.4.a. The Idea of Reconciliation.....</i>	<i>22</i>
<i>0.4.b. Reconciliation as 'Synthesis'.....</i>	<i>23</i>
0.5. Henry Sidgwick, and the 'Diversity of Methods'.....	25
<i>0.5.a. Sidgwick's Argument for Dual Reasonableness, and the Hope for Synthesis</i>	<i>28</i>

0.5.b. ‘The Truth Lies Between These Two Conclusions’: Sidgwick’s Moral Theory of Rational Benevolence.....	32
0.5.c. Reconciliation in Sidgwick: Rational Benevolence as a Synthesis Between Deontology and Utilitarianism.....	34
0.6. Sidgwick as a Deontological/Utilitarian Intuitionist, and the Personal Document.....	36
0.7. Sidgwick’s Place in History.....	41
0.7.a. Sidgwick at the Turn of the Twentieth Century.....	41
0.7.b. Sidgwick as a Utilitarian in the 20 th Century.....	44
0.7.c. Recognition of Sidgwick’s Reconciliation.....	46
0.8. Rational Benevolence in the Field of Modern Moral Philosophy.....	48
0.8.a. Rational Benevolence and Moral Ambivalence: Descriptive and Normative Value.....	49
0.8.b. Limits to Rational Benevolence as a Moral Theory.....	51
0.8.b.i. Practical Limitations.....	52
0.8.b.ii. Theoretical Limitations: Rational Benevolence, Conflicts Between ‘oughts’, and R.M. Hare.....	53
0.8.b.iii. Theoretical Limitations: Other Rival Theories.....	60
0.8.b.iii.a) W.D. Ross: Deontological Intuitionism.....	61
0.8.b.iii.b) William Frankena: ‘Mixed Deontological Theory’.....	64
0.8.b.iii.c) Robert Audi: Ethical Intuitionism.....	66
0.8.b.iii.d) Derek Parfit: Triple Theory.....	67
0.8.c. Other Considerations on Rational Benevolence as a Moral Theory.....	71
0.9. Summary of Aims, and Structure of the Thesis.....	74
1. LITERATURE REVIEW: OTHER READINGS AND INTERPRETATIONS OF SIDGWICK.....	79
1.1. The Methods of Ethics: Reception and the Early Years.....	80
1.1.a. Initial Reactions.....	80
1.1.b. Sidgwick’s Utilitarianism and British Idealism.....	83
1.1.c. Sidgwick at the Turn of the 20 th Century: G.E. Moore and Bertrand Russell.....	87
1.2. Sidgwick in the Mid-Twentieth Century.....	92

1.2.a. <i>Sidgwick and W.D. Ross</i>	92
1.2.b. <i>Sidgwick in Act and Rule Utilitarianism</i>	94
1.2.c. <i>Sidgwick and William Frankena</i>	97
1.2.d. <i>Sidgwick and John Rawls: Sidgwick's Appearance in Political Deontology</i>	99
1.2.e. <i>Sidgwick Revitalised: Sidgwick and Derek Parfit</i>	104
1.3. Perceptions and Oversights of Sidgwick's Reconciliation Between Deontological and Utilitarian Principles	108
1.3.a. <i>Marcus Singer: An Early Rejection of Sidgwick's Reconciliation</i>	109
1.3.b. <i>Janice Daurio: The Methods as One Single Moral Theory</i>	112
1.3.c. <i>Brad Hooker: Synthesis as Coherentism in Sidgwick's Work</i>	115
1.3.d. <i>Antony Skelton: The Rejection of Sidgwick's Coherentism, and the Question of Reconciliation</i>	119
1.3.e. <i>Robert Audi, and the Idea of a Sidgwickian/Rossian Synthesis</i>	123
1.3.f. <i>Mariko Nakano-Okuno: Sidgwick and 'The So-Called Discrepancies Between Utilitarian and Kantian Ethics'</i>	126
1.4. Current Wider Readings and Interpretations of Sidgwick	129
1.4.a. <i>Sidgwick and Peter Singer: The Point of View of the Universe</i>	130
1.4.b. <i>Sidgwick and David Phillips: Intuitionistic Utilitarianism, and The Unfairness Objection</i>	140
1.4.b.i. <i>'Intuitionistic Utilitarianism'</i>	141
1.4.b. ii. <i>'The Unfairness Objection', and the Question of Self-Evident Principles</i>	145
1.4.c. <i>Thomas Hurka: Sidgwick's 'Extreme' Deontology</i>	152
1.4.d. <i>Roger Crisp: Sidgwick's Deontology, and the Lack of Rational Synthesis</i>	156
1.5. Conclusions	160

2. BUILDING A SYNTHESIS: SIDGWICK'S ACCOUNT OF DEONTOLOGY	163
2.1. Sidgwick's Deontology in I:I: Initial Connections with the Intuitional View	165
2.1.a. 'Good' and 'Duty' in <i>Right Conduct</i>	165
2.1.b. Further Outlines of 'Intuitionism' in I:I	167
2.2. Reason, 'Ought', and the Deontological Version of Duty in Sidgwick	168
2.2.a. Reason, Right, and 'Ought': Sidgwick's Appeal to Kant	170
2.2.b. The Deontological Basis of Sidgwick's Theory of Moral Obligation	172
2.2.c. Consequentialism in Sidgwick's Theory of Moral Obligation	175
2.3. Sidgwick's Expansion on Deontology as it is Commonly Held	177
2.3.a. The Place of Common Sense Morality	177
2.3.b. The Question of Self-Evident Principles: Sidgwick's Critique of Intuitionism	181
2.4. The Deontological Nature of Justice, Prudence, and Universal Benevolence	186
2.4.a. Philosophical Intuitionism	187
2.4.b. Justice and Prudence as Deontological Principles	189
2.4.c. Universal Benevolence as a Deontological Principle	192
2.4.d. Rational Benevolence as a Deontological Duty	194
2.5. Sidgwick's Intuitionism, and Sidgwick's intuitionism	196
2.6. Conclusions on Sidgwick's Deontology	198
3. BUILDING A SYNTHESIS: SIDGWICK'S ACCOUNT OF UTILITARIANISM	202
3.1. Utilitarianism in I:I: Right, Good, and Happiness	205
3.1.a. The Roles of 'Right' and 'Good' in I:I	205
3.1.b. Rightness in I:I: Happiness as a Rational Ultimate End	207
3.2. The Rejection of Right as 'fitness to some end'	210
3.2.a. The Rejection of Right as 'fitness to some end': The Challenge to the Principle of Utility	212

3.2.b. <i>The Rejection of Right as ‘fitness to some end’: Sidgwick’s Teleology</i>	214
3.2.c. <i>Conclusions on Sidgwick’s Definition of Rightness</i>	217
3.3. Happiness and Pleasure as Ultimate Ends	218
3.3.a. <i>Sidgwick’s Rejection of Pleasure</i>	219
3.3.b. <i>Sidgwick’s Interpretation of Happiness as an Ultimate End</i>	222
3.4. Sidgwick’s Utilitarianism	225
3.4.a. <i>‘Difficult Calculations’</i>	227
3.4.b. <i>Utilitarianism as the Supreme Moral Principle? Positive and Negative Proof</i>	231
3.5. Conclusions on Sidgwick’s Utilitarianism: The ‘friendly alliance’	237
 4. ‘RATIONAL BENEVOLENCE’: SIDGWICK’S ACCOUNT OF THE SYNTHESIS BETWEEN DEONTOLOGICAL AND UTILITARIAN PRINCIPLES	240
4.1. The Meaning of ‘Synthesis’	242
4.2. The Possibility of Synthesis in Sidgwick’s Work	244
4.2.a. <i>Sidgwick’s Intended Ethical Neutrality</i>	244
4.2.b. <i>The Ideas of Reconciliation and Synthesis in I:I</i>	246
4.3. The Tone Changes: Enquiry vs. Conclusions, and the Dissolution of Sidgwick’s Neutrality	250
4.3.a. <i>I:III – I:VIII – The Early Influence of Reason on Sidgwick’s Neutrality</i>	251
4.3.b. <i>Sidgwick’s Abandonment of Neutrality in Book III</i>	254
4.4. The Synthesis: Sidgwick’s Moral Theory of Rational Benevolence	257
4.4.a. <i>Reiteration of Chapters 2 and 3</i>	258
4.4.b. <i>The Maxim of Rational Benevolence: Sidgwick’s Moral Theory</i>	259
4.4.c. <i>The Deontology of Rational Benevolence: Deriving Utilitarianism</i>	262
4.4.d. <i>The Utilitarianism of Rational Benevolence: Contextualising Deontology</i>	267

4.4.e. <i>Rational Benevolence as the ‘Proof’ of Utilitarianism: Equal Necessity of Deontological and Utilitarian Principles</i>	270
4.4.e.i. <i>Rational Benevolence as Proof of Utilitarianism’s Rightness</i>	270
4.4.e.ii. <i>Negative and/or Positive Proof</i>	272
4.4.f. <i>The ‘Truth Between These Two Conclusions’: Deontological and Utilitarian Principles as Mutually Informing in Rational Benevolence</i>	278
4.5. Rational Benevolence as a Synthesis	281
4.6. Sidgwick’s Synthesis, in His Words: ‘Professor Sidgwick’s Account of the Development of His Ethical View’	284
5. CONCLUSIONS	290
5.1. Restatement of the Problem	290
5.2. Restatement of the Solution: Henry Sidgwick’s Rational Benevolence	292
5.3. Using Rational Benevolence as a Moral Theory	295
5.3.a. <i>The Descriptive Value of Rational Benevolence</i>	296
5.3.b. <i>The Normative Value of Rational Benevolence</i>	299
5.3.b.i. <i>Rational Benevolence as Validation of Moral Action</i>	300
5.3.b.ii. <i>‘Rationally Benevolent’ Actions</i>	303
5.3.b.iii. <i>Rationally Benevolent Actions and Universalizability</i>	304
5.3.b.iv. <i>Rational Benevolence as Limitation to Moral Action</i>	307
5.4. Conclusions on Rational Benevolence as a Moral Theory	310
5.4.a. <i>Potential Applications of Rational Benevolence: Prenatal Technologies, Autonomy, and Political Theory</i>	312
5.4.b. <i>Limitations to Rational Benevolence</i>	317
5.4.c. <i>Final Comments</i>	321
BIBLIOGRAPHY	323

THE PROBLEM OF MORAL AMBIVALENCE

REVISITING HENRY SIDGWICK'S THEORY OF RATIONAL BENEVOLENCE AS A BASIS FOR MORAL REASONING, WITH REFERENCE TO PRENATAL ETHICAL DILEMMAS

INTRODUCTION

0.1. The Problem of Moral Ambivalence: 'The Parental Predicament'

In a hospital in "Exampleland", twelve pregnant women are meeting with their doctors¹. The women/families are not connected to each other in any way, except on two criteria. The first is that this pregnancy is, before any other considerations, a *wanted* pregnancy. The second is that, given a certain situation in which they find themselves, each of them will be making a decision as regards the future of that pregnancy. The thesis begins from the difficulties of these decisions, and the observation of the divergence in chosen courses of action, which is demonstrated by the following six possible scenarios:

- Parents A1: Receive a positive result from a prenatal test at an early stage of the pregnancy that indicates the child will be severely disabled; they consider their options, and decide to terminate [the pregnancy].
- Parents A2: Receive the same result at the same stage of pregnancy; they consider their options, and decide to continue [the pregnancy].

¹ Couples are usually assumed in the examples here, purely for ease of reference, but it is not necessary to the example that all of the pregnancies include both biological parents. The mother may be single; it may be a same sex relationship; it may be a situation in which only one or neither is the biological parent. In fact, all of these variations may well be directly relevant to the particular circumstances of that pregnancy. One criterion that is important however is that this is an actual *in utero* pregnancy, in which the/a mother has conceived. The thesis does not necessarily exclude the concept of pre-implantation embryonic screening, and this context will be briefly addressed in light of the thesis' findings in the conclusion - but this practice does not generally involve the same moral intensity as actual pregnancies, and is therefore not as relevant to the moral ambivalence that is central to my arguments here.

- Parents B1: Receive a positive result from a prenatal test at an advanced stage of the pregnancy that indicates that the child will be severely disabled; they consider their options, and decide to terminate.
- Parents B2: Are in the same situation as B1; they consider their options, and decide to continue.
- Parents C1: Have a healthy pregnancy, but they already have a child who is seriously disabled and requires a very high level of care; they consider their options and decide to terminate.
- Parents C2: Are in the same situation as C1; they consider their options and decide to continue.
- Parents D1: Are informed that the foetus itself is healthy, but the physical or mental health of the mother is at risk; they consider their options, and decide to terminate.
- Parents D2: Are in the same situation as D1; they consider their options, and decide to continue.
- Parents E1: Are informed that the foetus is healthy, but they are extremely poor, and already have four other children; they consider their options, and decide to terminate.
- Parents E2: Are in the same situation as E1; they consider their options, and decide to continue.
- Parents F1: Both the parents are fourteen years old. They consider their options, and decide to terminate the pregnancy.
- Parents F2: Are also fourteen years old. They consider their options, and decide to continue with the pregnancy.

These are only a few examples of the sorts of difficult situations in which parents could find themselves during a pregnancy. Into this equation we can insert an almost infinite amount of variations, and different problematic circumstances and factors that could bring the continuation of a pregnancy into question². The point of the examples

² The father may be seriously ill, for example, or the parent(s) may have other dependents such as elderly parents, and feel unable to also have a child while in that situation. There is also the question of *degrees* of severity of a situation – the degree of severity of a prenatally diagnosed disability (which sometimes cannot be detected, as is the case with Down's Syndrome), or degrees of poverty, or degrees of the mother's ill health. See also the Reproductive Review, 22nd April 2015, which published the statistic that in 2014, 70 women had abortions in the UK after suffering Hyperemesis

however is to draw attention to what they do all share in common. This is a grave ambivalence regarding the future of the pregnancy³. Each pregnancy is (in these examples) theoretically wanted, but in each case there are also *good grounds* on which to consider whether or not to actually continue. Defined in plain terms the two positions can be summarised as on the one hand a desire for the life of that pregnancy, and on the other a concern for the welfare of one or more of the parties involved - or for the *quality* of the lives involved. The resulting ambivalence between the two is a highly complex blend of emotional and ethical considerations, in which each course of action has different values, different qualities, and different bases for justification, and in which these values, qualities and justifications for *both* appear

Gravidarum, a serious type of morning sickness. Although a few of the pregnancies were unplanned, the majority were wanted.

³ There are clearly many individuals who will be excluded from these examples on the basis that they do have a strong inclination towards one course of action or another. The grounds for this are often religious, although not exclusively; even outside of religious circles, there are fierce debates over the moral status of the embryo, many believing life to begin either at conception, or at least at the point in the first trimester of pregnancy at which the embryo is thought to have developed enough to be able to feel pain (see Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press: 1993), 135-174; Michael Tooley, 'Personhood', in, Peter Singer and Kuhse (eds.) *Companion to Bioethics* (Oxford: Blackwell Publishing: 2009), p. 129). At the other end of the spectrum, many individuals will feel that they are responsible for not bringing a life into the world that is likely to have little potential for flourishing, or freedom from pain and suffering. (This responsibility may be felt privately, as in the majority of cases in which parents want to avoid bringing this sort of distress to their own children/families, or more publicly, by those who believe that now the technology that allows this avoidance is available, there is actually a wider social moral obligation to use it (see John Harris, 'Clones, Genes, and Immortality', in, *Wonderwoman and Superman* (Oxford: Oxford University Press: 1998), p. 328; John Harris, *Enhancing Evolution; The Ethical Case for Making Better People* (Princeton: Princeton University Press: 2007) ; John Harris, 'Enhancements are a Moral Obligation', in, Nick Bostrom and Julian Savulescu (eds.), *Human Enhancement* (Oxford: Oxford University Press, 2011) p. 131 ; Julian Savulescu, 'Why We Should Select The Best Children', in, *Bioethics*, Vol. 15, No.5/6 (2001) 413-426). Then there is the position of the disabled community's critique on abortion, which holds that to terminate pregnancies on the grounds of disability is discriminatory, and simply a form of disguised eugenics. See Tom Shakespeare and Anne Kerr (eds.) *Genetic Politics: From Eugenics to Genome* (Cheltenham: New Clarion Press, 2002); Daniel Kevles, *In the Name of Eugenics; Genetics and the Uses of Human Heredity* (Cambridge, MA: Harvard University Press, 1995); David Galton, *Eugenics; The Future of Human Life in the 21st Century* (London: Abacus, 2001); Jonathon Glover, *What Sort of People Should There Be?* (London: Penguin Books, 1984), and, *Choosing Children; Genes, Disability, and Design* (Oxford: Clarendon Press, 2006); Phillip Kitcher, *The Lives to Come: The Genetic Revolution and Human Possibilities* (New York: Simon and Schuster, 2001); Nick Agar, *Liberal Eugenics: In Defence of Human Enhancement* (Oxford MA: Blackwell, 2004).

In short there may be a myriad of different reasons why parents are *not* ambivalent, and the vast majority of these are justifiable, as the thesis will indirectly show, through the argument that *both* courses of action can be found to be right simultaneously. All that is important here is that as they do not experience the ambivalence which I am using to demonstrate the point that both courses of action can be found to be right simultaneously, they are not included in the sorts of examples listed above, and therefore are not directly relevant to the thesis.

equally relevant to the situation⁴. In scenarios that are otherwise identical, two fundamentally different courses of action appear to emerge as the 'right' thing to do.

In the context of prenatal decision making, there is a further contributing factor to this already very sensitive situation as it is experienced by individuals at a private level. This is that both the continuation and the termination of difficult pregnancies are independently supported – often very vocally - at the public socio-political level⁵. Although the institution of individual autonomy is designed to protect parents'

⁴ Clearly, this ambivalence is both emotional and moral. It is the personal and subjective emotional sense that both courses of action are somehow right that informs the position of moral ambivalence, but any objective moral sense that both courses of action are right will in turn intensify the emotional uncertainty. Given the moral philosophic character of the thesis however, this emotional and ethical ambivalence will generally be referred to as 'moral ambivalence'.

⁵ Societal pressure on individual ethical decisions is not limited to the issue of abortion, but it is particularly noticeable – and the opinions particularly strong – in this context. Parents frequently report that they encounter the influence and weight of other viewpoints while making their own personal decisions. In an interview with Jane Fisher, Chief Executive of British charity organisation Antenatal Results and Choices (ARC), 12th December 2013, Ms Fisher confirmed to me that parents who seek the services of ARC often feel the expectations of society for them to choose either this course of action or that, and the subsequent sense of judgements, even more acutely than they feel their own beliefs on the matter. Ms Fisher also spoke about how the ARC building is often targeted by protestors from the pro-life movement, who bring babies to their protests, as if 'somehow implying that we don't know what newborn babies look like'. This has frequently caused great distress to pregnant women who are at ARC at the time, using their services to decide on whether or not to continue their pregnancies. Ms Fisher also referred to what she terms as the 'disney-fication' of certain genetic conditions that were once held to be serious but that now - with better treatments, better social integration between disabled and non-disabled communities, and the societal de-stigmatising of certain conditions - society generally considers to no longer be good grounds for abortion. The most obvious example is Down's Syndrome: Many individuals with Down's Syndrome lead independent, fulfilling and largely normal lives, and society recognises this. However, the severity cannot be predicted *in utero*, and individuals with Down's Syndrome may have serious heart conditions, and/or very limited capacity for learning or independence, and many other health problems. The 'disney-fication', that Ms Fisher refers to, is the failure of this more 'tolerant' society to recognise that Down's Syndrome is still for many a debilitating, life limiting and distressing condition. On the other side of the debate, Ms Fisher also reported that many parents have admitted to feeling judged by their friends or family for continuing difficult pregnancies where the health or life of the future child is in question, when it appears as though they are simply willingly bringing a child into the world to suffer. They are told that they could 'easily try again', that they could/should avoid the trauma, or, most worryingly, that it would avoid great financial costs to the parents, and even to health services.

There is also evidence to suggest that some healthcare professionals in this area, especially genetic counsellors who advise on possible outcomes where a genetic condition is concerned, are not acting impartially, and are in fact advancing their own views to parents on the course of action they should take (see Merryn Ekberg, 'Maximising the Benefits and Minimising the Risks Associated with Prenatal Genetic Testing', in, *Health, Risk, and Society*, Vol.9, No.1 (2007), 67-81, and, 'Second Opinions: The Old Eugenics and the New Genetics Compared', in, *Social History of Medicine*, Vol.20, No.3 (2007) 581-593; Tony Fitzpatrick, 'Before the Cradle: New Genetics, Biopolicy, and Regulated Eugenics', in, *Journal of Social Policy*, Vol.30, No.4 (2001) 589-612; Tom Shakespeare, 'Losing the Plot'? Medical and Activist Discourses of Contemporary Genetics and Disability, in, *Sociology of Health and Illness*, Vol.21, No.5 (1999) 669-688; Amos Shapria, 'Wrongful Life' Lawsuits for Faulty Genetic Counselling', in, *Journal of Medical Ethics*, Vol.24, No.6 (1998) 369-375; Christy Roberts, Laura Stough, and Linda Parrish, 'The Role of Genetic Counselling in the Elective Termination of Pregnancies Involving Disabilities', in, *The Journal of Special Education*, Vol.36, No.1 (2002) 48-55). This seems to occur particularly when the counsellor believes the 'best' course of action would be to end the pregnancy.

decisions from interference by others, very rarely do we make our decisions in a social vacuum, and any pressure or stigmatisation that parents encounter from either position only serves to intensify the difficulties of which they are already aware⁶.

Taken overall, I refer to this contribution of private and social factors to the serious ambivalence that can arise between two justifiable but fundamentally opposing courses of actions regarding the future of a pregnancy, as the 'Parental Predicament'. The context of the Parental Predicament provides, I believe, one of the most compelling examples of human moral ambivalence – and it is the *cause* of this ambivalence that provides the thesis with its foundational claims about morality from which its main aim is developed. This cause of ambivalence is now examined in the following section.

0.2. The Significance of Moral Ambivalence: Two Types of Moral Consideration as 'Reasonable'

The word 'ambivalence', as representative of the ethical and emotional state of affairs in the Parental Predicament, has been chosen purposefully here. To be ambivalent does not, contra to some popular usage, mean to have little or no opinion either way, but rather to hold two strong but different opinions at the same time - as is demonstrated by the divergent decisions made in the Parental Predicament. In broader moral terms, that divergence translates into two types of moral consideration. The first – the desire to continue the pregnancy – reflects a recognition that there are certain fixed principles that are felt on some intuitive or instinctive level to be inviolable and universal, and as such to be 'right' in an absolute sense. In this case the inviolable or absolute value is that of human life, or a potential

⁶ In terms of evidence that parents do feel this way, there is an extensive amount of academic and clinical literature. See Paivi Santalahti, Elina Hemminki, Anne-Marie Latikka, and Marrku Ryyanen, 'Women's Decision-making in Prenatal Screening', in *Social Science and Medicine*, Vol. 46, No.8 (1998) 1067-1076; Issac Rabino, 'Genetic Testing and its Implications: Human Genetics Researchers Grapple with Ethical Issues', in *Science, Technology and Human Values*, Vol. 28, No.3 (2003) 365-402; M.H. Parker, K.L. Forbes and I.Finlay, 'Eugenics or Empowered Choice; Community Issues Arising from Prenatal Testing', in *Australian and N. Zealand Journal of Obstetrics and Gynaecology*, Vol. 42, No.1 (2003) 10-4; and V. Seavilleklein, 'Challenging the Rhetoric of Choice in Prenatal Screening', in *Bioethics*, Vol. 23, No.1 (2009) 68-77, as just some examples of such literature. Also see below for my argument that autonomy is an incomplete solution to the problem of moral ambivalence, precisely because of its function in allowing for freedom of opinion and speech.

human life⁷. The second position – which demonstrates concern for the welfare or quality of life of the relevant parties – reflects the inclination to consider what the outcome, or the consequences, of an action might be, and to decide on the course of action accordingly.

Both of these viewpoints are, as noted above, equally valid. But it is my view that it is actually the space between the two apparently justifiable positions that is far more revealing of the reality of the human moral experience than positions that defend exclusively either one course of action or another. This is because, I argue, the two positions are simultaneously valid because they are both simultaneously *reasonable*. By ‘reasonable’ here I mean of sound judgement, objectively or impartially justifiable, and that the principle or course of action is intrinsically defensible by appeal to nothing but itself, i.e. it is not contingent on anything further for its justification⁸. Principles that are reasonable in this way are, in short, unconditional. Applying this definition to the two positions in the Parental Predicament, a decision to continue a difficult pregnancy is reasonable because the desire to continue and preserve a human life has intrinsic, or unconditional, value in itself, without needing to refer to further reasons for that value⁹. An eventual decision to end a difficult pregnancy is reasonable because the action of avoiding bringing pain, suffering, or a threat to welfare of any kind also has intrinsic and unconditional value of the same kind. The fact that both positions share this kind of reasonableness is, I argue, the actual *cause* of moral ambivalence, and the deeper

⁷ It does not actually matter here whether the parents believe the unborn child to represent a whole human life at this stage, or only a potential one; it only matters that they are ambivalent about the weight of this absolute value when compared to the weight of outcome-based values. In this way, the thesis avoids the problematic ‘beginning of life’ moral debate, in which it is disputed exactly what is to count as ‘human life’, and where – in a stage of pregnancy or even pre-implantation – this is thought to become applicable. Parents may even believe that life does begin at conception, and still be ambivalent regarding their course of action in a very difficult pregnancy. Equally, parents might believe that ‘life’ begins at a later stage in pregnancy, and yet still be ambivalent regarding the earlier stages. I emphasise here the vast scale of positions, and the variable degrees within those positions, that could constitute the situation for individuals in the Parental Predicament.

⁸ Aside from this cursory definition of reasonableness, I do not mean to presuppose here a formal, ultimate definition of what is ‘reasonable’. It is actually part of the thesis’ remit to argue for a wider understanding of what is morally reasonable. Nor do I mean to assume here that I am right that both of these positions are reasonable. Many would obviously disagree with that insinuation. However, I am working from the point of view of moral ambivalence, and it is my view that this ambivalence is caused by the equal reasonableness of two types of moral principle – the thesis goes forwards on this premise, which I will aim to justify it via appeal to Sidgwick’s understanding of the relationship between ‘reasonable’, and ‘right’.

⁹ See Samuel Scheffler, ‘Morality and Reasonable Partiality’, in, Brian Feltham and John Cottingham (eds.) *Partiality and Impartiality: Morality, Special Relationships, and the Wider World* (Oxford: Oxford University Press: 2010)

explanation of why both courses of action appear to be valid, or why they both appear to be *right*¹⁰. Because both types of moral consideration can be described as reasonable in this way i.e. as not needing reference to anything further for their validation, both naturally form an integral part of our moral processes – as indeed they are seen to do, in grave ethical dilemmas such as the Parental Predicament. ‘Ambivalence’ thus represents the fundamental premise of the thesis, which is this observation that in situations of great emotional and moral intensity, that intensity is often caused by the conflict not just between two apparently valid courses of action, but between two courses of action – or two types of moral consideration - that are both defensible on the basis that they are both reasonable, both in some way *right*, and that are both relevant to the human moral experience.

To define moral ambivalence as the situation in which two positions are simultaneously reasonable has another implication, aside from supporting the view that both types of moral consideration – absolute-principle based and outcome based – are right¹¹. This must also be addressed in order to gain a full understanding of what is meant when it is said that both positions are reasonable. It is the paradoxical but necessary inference that if each position is right, and yet they are fundamentally opposed to each other, then each position must also be, to some extent, wrong. It would seem impossible to maintain that life has an absolute value, and also to maintain at the same time that life depends upon certain quality-based

¹⁰ ‘Right’ here is intended mainly to represent how the parents might describe the situation. I do, however, mean to suggest that there is a certain truth, or certainty, or fact in both of those positions, as demonstrated by their respective reasonableness, and that this does indicate a certain rightness. This implies that I am adopting the position in which reason and ‘rightness’ are related. As this was Henry Sidgwick’s position, it is also mine. I am aware however that the use of the word ‘truth’ as representative of what is meant by ‘reasonable’, and therefore ‘right’, is more problematic when it comes to the matter of whether there is a link between ‘truth’ and moral justification. Jurgen Habermas famously argues that truth and moral rightness are not the same thing (Habermas, *Truth and Justification* (Cambridge MA: MIT Press: 2003). Steven Levine, however, agrees that truth and moral rightness can be linked (Levine, ‘Truth and Moral Validity: On Habermas’ Domesticated Pragmatism’, in, *Constellations*, Vol.18, No.2 (2011) 244-259). I am actually not arguing specifically for that link. I use ‘truth’ here as a superficial way of describing what I mean by reasonable, but ultimately I argue – via Sidgwick - that it is only principles that have been established via a process of reasoning that can be called morally right.

¹¹ By ‘absolute-principle’ I am referring to the ‘fixed principle’ position, which holds that certain principles such as ‘do not kill’, are always right; I do not mean that only these principles are unconditional, because as I have been arguing – and as I will argue overall in the thesis - it is my view that certain outcomes (such as the avoidance of suffering) are also unconditional. The terms ‘absolute principles, or ‘absolute value’ are simply used to denote the moral idea that certain *rules* are always right, as distinguished from the moral idea that certain *outcomes* are right – as despite their shared unconditionality they do represent different ethical approaches, and in order to understand this to be the cause of moral ambivalence we require a way of identifying the two positions.

conditions, such as freedom from pain, or the ability to flourish. The two types of moral principle therefore appear to be mutually exclusive, with each reasonable position also incurring wrongness precisely from the reasonableness of the other. To act on the values of one course of action is, essentially, to ignore the moral values of the alternative.

At a purely practical level of the Parental Predicament, the two positions are of course mutually exclusive; either one or the other will be decided upon, and there is no 'middle ground' in this situation¹². However, the overall crucial point that I wish to maintain here about moral ambivalence is that it is a fact – as is the existence of those two types of moral consideration, which conflict to create that ambivalence. It is very difficult to deny that both types of moral consideration do appear in our moral thought processes; simple introspection on the part of most people will reveal inclinations towards both absolute principles and outcome based moral principles. Furthermore, it also seems undeniable that both are important, that both types of moral consideration carry significant and persuasive values - and that they often appear together, in relation to the same issue¹³. I argue overall then that both types of moral consideration are at once simultaneously reasonable – and incontrovertible.

0.3. Moral Ambivalence, and the Problem of the 'Divide' In Moral Philosophy

It was my aim in the previous section to argue that as both types of moral consideration are reasonable, they are both an inevitable part of our moral processes – and that this is both logically and empirically the case. But, I argue, there is a major obstacle to the acceptance of this understanding of human morality,

¹² The two positions might not be mutually exclusive in all contexts however – and it may be possible in some situations to find a way in which the values of each position are both accounted for in a moral judgement. This will be addressed in the conclusion, in light of the thesis' findings.

¹³ I suggest a thought experiment here, that I believe supports my claim that both absolute values and outcome-based values are often active in the process of making a moral judgement *at the same time*. This experiment involves imaging a scenario in which a motorist is driving on a motorway when a pedestrian appears in front of them. I theorise that the first impulse of the motorist would be to swerve to avoid hitting the pedestrian, because the initial instinct is to avoid killing someone, which is an absolute-value based response. However, I suggest that the next – and almost *simultaneous* – impulse would be the realisation that swerving into the neighbouring lane would very likely cause a huge accident, which would almost certainly kill more people, and that this is an outcome-based value response. As this 'pedestrian on the motorway' scenario involves the fact that the motorist/thinker is travelling at speed, and would therefore have little time for prolonged thought or deliberation, it is intended to demonstrate a) just how quickly the two moral impulses can appear alongside each other, and b) thus the natural co-existence of these different but equally valid and important moral inclinations.

and it occurs in moral philosophy itself. In the following section I outline this problem as it occurs in moral philosophy, and argue that it is synonymous in its cause with moral ambivalence as presented above. I then contextualise the problem in terms of the failure of individual autonomy to provide a solution to the ambivalence of the Parental Predicament.

0.3.a. The ‘Divide’ in Moral Philosophy: Ambivalence Between Theories

Within moral philosophy there are two dominant interpretations of how to define what is moral¹⁴. These are a) the deontological school of thought, which originated mainly with Immanuel Kant in the 18th Century and which maintains that moral value lies only in the performance of categorically prescribed, unconditional, and non-consequentialist moral duties, and b) the utilitarian school of thought, introduced in its most recognisable form through the work of Jeremy Bentham and J.S. Mill, which argues that moral rightness is to be found not in actions themselves per se, but in the attainment of certain non-moral ends, such as happiness, or other ‘goods’¹⁵. Each school of thought has a long and varied tradition in which it has experienced support and decline, revisions and additions – and both are essential to any understanding of the moral philosophy of the 20th Century, and of the modern day. A

¹⁴ They are not the only interpretations, but I argue that they are the most important/relevant to my arguments here.

¹⁵ These are very broad interpretations of those terms. It is especially possible for confusion to arise over the term ‘deontology’, as technically ‘*deon*’, or ‘duty’, could be used in any system of ethics in which one adheres to duties or rules of some kind – and these rules could be of a utilitarian kind. (An example of this very point is provided by father of utilitarianism Jeremy Bentham, whose seminal work of moral philosophy was actually entitled ‘*Deontology*’, on the basis that he perceived the securing of happiness to be the ultimate moral duty). For this reason, I have mainly avoided referring to the deontological position as ‘duty’ based. For the purposes of this thesis, the terms deontology and utilitarianism are understood in the following ways: Deontology is taken to refer to school of thought (associated mainly with Kant) in which morality lies in certain fixed, absolute principles held to be categorically right without reference to any consequences. As these are most traditionally called ‘duties’ as per the term ‘deontology’, where I do use the language of duty, it is this deontological understanding that is meant. Utilitarianism is taken to be the system of ethics that places moral value in non-moral ends, such as Happiness, or good, and in which actions are right or wrong in as far as they do or don’t produce those outcomes, or consequences. ‘Classical utilitarianism’ will refer to the doctrine as it appears under Bentham, Mill (and occasionally Sidgwick, although I dispute that Sidgwick can be considered a classical utilitarian, as I will argue) – the first definition I gave may be called ‘traditional’, or ‘standard’ utilitarianism’, and is intended to embody the typical features of utilitarianism as it generally appears in all other utilitarian theories since Bentham and Mill. For consistency, these two moral positions will now generally be referred to as ‘absolute-principle based’ and ‘outcome based’ respectively.

serious problem, however, lies in the relationship between them. Deontology and utilitarianism clearly represent positions that are fundamentally incompatible; the placing of intrinsic moral value in certain categorically prescribed actions themselves is *prima facie* irreconcilable with the argument that moral value lies in non-moral ends, and in the conditional actions that secure those ends. Yet both moral schools lay claim to being able to explain the truth of human morality, and in these claims they are both robustly supported, by the decades' worth of moral theorists who have defended them. That is, deontology and utilitarianism appear to present two valid interpretations of morality, that are also mutually exclusive – and the relevant moral philosophy has thus presented the two positions as being fundamentally divided¹⁶.

This divide, it will be noted, is an exact model of the very situation just outlined in the previous section - that situation in which there seems to be within the human moral experience two equally reasonable moral impulses that are fundamentally at odds with each other. Translated into the direct terms of the Parental Predicament (and many other ethical conflicts), the view that human life has intrinsic and unconditional value represents the fixed or absolute-principle deontological position; the view that it is intrinsically and unconditionally right to avoid outcomes that will involve pain, suffering, and threat to welfare represents the utilitarian position¹⁷. It is clear then that the 'divide' between deontological and utilitarian understandings of morality at the philosophical level directly reflects the divide between competing reasonable moral considerations as they occur in the lived moral experience; the divide is, essentially, another form of moral ambivalence.

0.3.b. The Problem of the 'Divide': Autonomy and the 'Illusion of Choice'

¹⁶ An outline of this 'divide' appears in the literature review. Although the purpose there is to examine how and where Sidgwick's work has been used throughout the years, the thread of the divide between deontological and utilitarian positions naturally emerges as well, as this is the framework of moral debate throughout which the place of Sidgwick's work is traced.

There have, over the years, been theories that have aimed to unite deontological and utilitarian considerations, and some have had moderate success, in some ways. As these theories are a part of the field to which this thesis is directly relevant, they are outlined below.

¹⁷ I argue that this conflict does not just pertain to the Parental Predicament, but to the majority of ethical dilemmas. So many answers to the moral question of 'what ought I to *do*', fall either into the category of obeying universal and absolute moral principles, or the category of acting in such a way that the best outcome will be secured.

The main practical solution to the prevailing conflict between absolute-principle based and outcome based moral considerations has been to implement freedom of choice. The Parental Predicament is again the best example here. Current healthcare policy in this area aims to manage the divergence between equally defensible courses of action by emphasising individual autonomy, the socio-political framework which defends the right of individuals to make their own choices while other individuals are also allowed to make theirs¹⁸.

At first glance, individual autonomy gives the impression of being a highly effective tool for a situation in which there are two alternative positions on the same matter. But something crucial is missing from autonomy as a solution to the problem of moral ambivalence. Firstly, autonomy has *nothing of substance* to say about the actual decisions themselves. What is overlooked by autonomy in this context – as vital as the concept is for a free society – is that precisely what is revealed by parental ambivalence is this fact that parents are morally and emotionally unsure of which decision to make. The personal internal struggle experienced by those parents is not in any helpful way addressed simply because the choice is theirs to make. Furthermore, autonomy itself can actually *cause* a problem in the context of serious ethical dilemmas. The dual positive and negative angles of liberty entrenched by autonomy mean that the liberty that allows parents the freedom to choose the course of action best for them is precisely the same liberty that allows the free expression of other opinions. For a context in which opposing views are often propounded so vociferously, this arrangement is hardly going to provide a solution. Social pressure has already been noted above as being part of the Parental Predicament – but framed in terms of autonomy, it becomes clear that in the absence of a moral philosophic means of understanding human morality to be comprised of both absolute-principle-value and outcome-value (or deontological and utilitarian) aspects, autonomy can give no meaningful practical guidance beyond simply defending each course of action separately. And this, ironically, is precisely part of what causes that problem in the first place.

¹⁸ Autonomy, by its very nature, has a dual function that appears to make it perfect for the task of defending private, personal choice. There is positive autonomy, which accounts for a person's 'freedom to', and negative autonomy, which accounts for the person's right to be 'free from' [interference from others].

Ultimately then, I argue, the choice provided by the framework of autonomy is little more than an illusion. *Because* both positions *are* valid, both positions *are* going to appear in autonomy, juxtaposed just as they are without it. The endless circular tautology in the Parental Predicament between how far each type of value is relevant and how far the limits of the alternative are relevant is directly reflected by that question inherent in the institution of autonomy, of how far individual choice should be protected and how far it should be endorsed¹⁹. For this reason, autonomy also directly embodies and reflects the ambivalence as it occurs at the moral philosophical level. Autonomy is undoubtedly important, and it should form a part of any ethical policy implementation, but it is a framework only, and not adequate by itself to provide a defence of the ethical decisions made within it. In plain terms, autonomy on its own simply leaves the door open to the perpetuation of moral ambivalence as it appears at a public level.

0.4. Challenging the Divide: Reconciliation and Synthesis Between Deontological and Utilitarian Principles of Morality

Given all the observations made above, the question must now turn to how best to address this lived and inter-theory moral ambivalence. I have argued that its cause is that both absolute and outcome (or deontological and utilitarian) considerations of morality are reasonable, in that they both account for valid ethical and emotional concerns. It has also been seen that these concerns appear to be fundamentally incompatible, and that this leads to situations of painful and protracted deliberation - in which individuals in the contexts such as the Parental Predicament are also left vulnerable to the effects of *public* ambivalence. I have argued further that there is, at the level of moral philosophy, a divide between deontological and utilitarian properties of morality that is directly symptomatic of the divide as it is seen in lived experience: There are two theories that defend two compelling arguments as to what is moral, and these arguments are *prima facie* divergent. Not only then is moral ambivalence as it occurs in lived experience difficult enough in itself, but the moral

¹⁹ Within political theory, again, autonomy itself can offer nothing. It is the precise role of autonomy to paradoxically both ensure limits and to remove them, to allow free choice but to check it. It justifies as much as it limits, and limits as much as it justifies.

theories that inform the various values only further compound the very problem that they aim to address. Whereas there have been some theories that have aimed to unite the two types of moral property to some extent, these have only been partly successful; the divide still pervades moral philosophy, and it is still affecting ethical theory in a variety of contexts²⁰.

I argue that this opposition between deontology and utilitarianism as it occurs at the moral theoretical level is erroneous. I have aimed to show in the above discussion on ambivalence that the two types of moral consideration are both fundamental, reasonable, and, as such, irrefutable aspects of human moral decision making, and I argue on this basis that to interpret morality to be exclusively one or the other is an inaccurate understanding of the reality of that morality. I therefore argue in the same line that the deontological and utilitarian understandings of morality that represent those moral considerations at the academic philosophical level are *also* both reasonable – and therefore also that to argue exclusively for one position or the other is an inaccurate understanding of the reality that they are trying to represent. The opposition between them is therefore also unhelpful as a means of informing important decision making frameworks such as autonomy, *and* of supporting the individual difficult ethical decisions being made therein. For as long as they are posited as separate views of morality, the endless oscillation between them will simply continue. As Henry Sidgwick pointed out, “...if there are different views of the ultimate reasonableness of conduct...it is easy to see that any single answer to the question ‘why’ will not be satisfactory, as it will be given only from one of these points of view, and will always leave room to ask the question from some other”²¹.

²⁰ Despite his own optimistic description of utilitarianism being the purest form of deontology, political philosopher Will Kymlicka draws attention to a distinction that *still* exists between deontological and utilitarian theories, and that has been made all the more palpable by the developments in ethical-political theory, such as an increase in global concerns, and the advances in health care. Kymlicka concludes that the question of whether the ‘right’ or the ‘good’ (thought to be represented by deontology and utilitarianism respectively) is to take priority is still ‘a central dividing point of contemporary political theories’. See Kymlicka, ‘Liberalism and Communitarianism’, in, *Canadian Journal of Philosophy*, Vol.18, No.2 (1998)181-203.

Alasdair MacIntyre, in a paper in 2013 entitled ‘*Having Survived Academic Moral Philosophy of the Twentieth Century*’ (in, Fran O’Rourke (ed) *What Happened in and to Moral Philosophy in the Twentieth Century* (Notre Dame, Indiana: Undpress: 2013) – the title of which is telling enough – also states that utilitarianism and Kantianism are, ‘with some rare and sophisticated exceptions’, still presented as ‘incompatible and rival standpoints’, in which adhering to one simply means rejecting the other.

²¹ Sidgwick, *The Methods of Ethics* (London: Macmillan: 1874) I:I:6

In the next two sections I give an overview of how I propose this dual reasonableness should be accounted for, and therefore how this divide should be addressed. This is done in preparation for the next half of the chapter, which contains the details of how this possible solution is to be drawn from the work of Henry Sidgwick.

0.4.a. The Idea of Reconciliation

It is my view that there is only one point from which to begin speculating as to what the solution to the theoretical ambivalence between deontology and utilitarianism might be – and that this is present in the very fact of moral ambivalence itself. It is, quite simply, the acceptance that both absolute-principle based and outcome based properties of morality *are* relevant to the human moral experience. To deny that either is important or relevant is, as I have argued above, wrong both factually and theoretically – and taking this position has clearly not been effective, either philosophically, or practically. It is for this reason that I have specifically been using the language of ‘division’ to refer to the state of affairs between deontology and utilitarianism. ‘Divide’ is intended to convey the idea that there was – and still is - in human morality actually a *whole* of which deontological and utilitarian principles are constituent parts. And if this is the case, the suggestion is that there is a means through which those parts, which have been so widely and for so long held in direct contention, can be *reconciled*.

It must be asked, however, what is meant by ‘reconciliation’, and why this notion is important when it would also be possible to simply present a better defence for each course of action separately. There is one answer to both of these questions - for by ‘reconciliation’ I do not mean that a better space is created for each position to simply co-exist independently of each other. Giving a better philosophical defence of termination or continuation of difficult pregnancies respectively could arguably give autonomy better resources with which to defend parents deciding on either course of action, but this would not improve the situation. Defending each course of action independently not only offers nothing new, but it fails to address the enduring tension between deontological and utilitarian moral theories by failing to address the reality of the human moral experience – that reality that is demonstrated by the dual reasonableness that causes ambivalence. It is this dual reasonableness that needs to be accounted for – and the more specific fact that both moral impulses are often

(in situations of grave moral ambivalence) present in our moral processes *simultaneously*. For what I argue is indicated by this simultaneous reasonableness is the fact that each type of moral property somehow *depends* to some extent on the values of the other. On this basis, I argue that the most effective way to account for this dual reasonableness of absolute-principle based and outcome based moral properties is not to defend the reasonableness of each independently, but to argue that they are actually in some way inherently connected. It is this point that is the most important for the concept of reconciliation as I am using it here, for it points to the more specific understanding of what I claim that reconciliation involves. This is the idea of there being - between deontological and utilitarian principles of morality – an actual *synthesis*, in which the two types of moral property are no longer seen as mutually exclusive, but rather as *mutually informing*.

0.4.b. Reconciliation as a ‘Synthesis’

My argument for there being a type of relationship between duty-based and outcome-based moral properties that could be called a ‘synthesis’ is drawn from the observation made about ambivalence above, that the simultaneous presence of two reasonable types of moral value must mean that the reasonableness of each type of moral consideration will place limits on the extent to which the other type is reasonable. I argue that this suggests that there is a fundamental connection between the two properties – a way in which they operate essentially as two parts of a larger whole. But this relationship is not just a negative one (which would eventually lead us back to the same stalemate of the current situation). Where the reasonableness of each moral value *limits* the other, the view could be reversed - as the situation of equal reasonableness was reversed to show how it leads to the limiting of each position - so that the relationship can be viewed positively, as one in which the reasonableness of each moral value actually *supports* the other. In such a relationship, the values of each type of consideration are not only crucial to the restriction of the other, but also to the *validity* of the other. In this way, they are truly mutually informing. The relationship is a dynamic and interdependent one, in which each moral property has value, a specific identity, and, crucially, a role in supporting and limiting the other. It is this understanding of the relationship that is intended by the use of the word ‘synthesis’. This is not merely a relationship of incidental co-

existence, but one in which components combine to form something else – in this case, a more inclusive and comprehensive understanding of the human moral experience. This is, I argue, the only way to hold both deontological and utilitarian principles in such a way that they are not continually in contention (as a mutually-limiting-only relationship would allow for), but are both seen to be vital and functioning parts of a whole that represents the reality of the human moral experience. Deontological and utilitarian properties would not be presented as simply being the *same* – each principle must remain independently valid in order to bring to the synthesis its own crucial moral properties. But where the nature of ambivalence suggests that neither deontological nor utilitarian moral principles are adequate by themselves to describe the full range of the human moral experience, a synthesis between them allows each to bring its own independent value, while itself only being fully realised by the value of the other. Within such a synthesis, the two types of moral property would thus remain different, but they would no longer be *divided*.

It is this understanding of morality, in which the traditional boundaries between deontological and utilitarian are largely dissolved, that would, I argue, offer to moral philosophy a more complete and accurate theory of human morality that is not constrained by either the values or the deficiencies of just one moral position. It is also a theory that is truly different to those that simply defend one course of action or another, and to those that have united utilitarian and deontological properties to some extent, but ultimately on the basis or terminology of either one position or the other. I theorise that a true synthesis between the two, like the one just described, would go a significantly further way towards ending the perpetual oscillation between them, and argue that this in turn could more effectively inform institutions responsible for preserving the necessary freedom of choice, such as individual autonomy²².

There is one last point to note about this idea of synthesis. This is that the argument for the mutually informing nature of deontology and utilitarianism might be accused of containing an obvious *saltrus in demonstrando*, in that just because two

²² Naturally, both courses of action would *also* be defended independently, even though this is not the main aim of the synthesis. But despite it not being the main aim, this is important too, given that ultimately in situations such as the Parental Predicament there will come a point at which either course of action will be taken - and there can be no 'middle-ground' between deciding to end a pregnancy, and deciding to continue it. As essential as I believe a synthesis between deontological and utilitarian moral concerns to be, it is, for that reason, just as important that those decisions are protected – and I do believe that the synthesis provides this protection, by default of its arguing for the reasonableness of both.

positions limit each other, this does not necessarily mean that they validate each other. But this mutual validation is precisely the aspect of the relationship between deontology and utilitarianism that I aim to develop in this thesis. I believe that this is possible, and that it is the relationship already demonstrated by the mutual limiting of the two types of value that strongly suggests it, as I argued above. Finally, I argue that this relationship for the mutual validation between deontological and utilitarian properties of morality can in fact be built on the basis of their shared reasonableness, and that the work of Henry Sidgwick provides this very model, in his theory of Rational Benevolence.

0.5. Henry Sidgwick, and the ‘Diversity of Methods’

In 1874, a Fellow of Trinity College Cambridge and lecturer in classics published a book in which he aimed to address the great variety of principles, codes, values, and beliefs that comprise the moral decision making of ordinary people. The book was *The Methods of Ethics*²³. Its author was Henry Sidgwick. Vast in scope and meticulous in its investigations, *The Methods* was instantly received as a classic²⁴. It is to this day recognised both as a masterpiece in its own right, and as one of the most seminal works in 19th and 20th Century western philosophy. At the time of its publication, Sidgwick was just thirty-six years old²⁵.

For as long as Sidgwick has been a presence in moral philosophy, he has almost unanimously been considered to be a utilitarian – a classical utilitarian, in fact, in the tradition of Bentham and Mill²⁶. Sidgwick’s work certainly followed theirs

²³ References to *The Methods* throughout the thesis are abbreviated to ‘ME’, and all references are to the 7th edition of *The Methods* (1907) unless otherwise indicated by a number appearing after ME. The only exception to this is where reference is made to ME’s various prefaces, where it will appear as ‘ME: x’, for example. Direct citation of ME appears in the volume’s own roman numeral and page number format, i.e. I:III:33.

²⁴ Even most of its critics admitted that it was exceptional for its depth, breadth, and attention to detail. See Hastings Rashdall ‘Professor Sidgwick’s Utilitarianism’, in *Mind* (1885) Vol.10, No.38 (1885) 200-226; James Seth, ‘The Ethical System of Henry Sidgwick’ in *Mind*, Vol.10, No.38 (1901) 172-187, and in particular C.D. Broad *Five Types of Ethical Theory* (Paterson, N.J: Littlefield Adams, 1930).

²⁵ *The Methods* was not Sidgwick’s only work, although it is certainly his magnum opus. Sidgwick was a political theorist and a social reformer, as well as an economist, an active member of debates on sexual politics and the rights of women, and a devotee of parapsychology. He wrote *Principles of Political Economy* (London: Macmillan and Co.) in 1883, *The Elements of Politics* (London: Macmillan and Co.) in 1891, and *Practical Ethics* (London: S. Sonnenschein) in 1898, among a huge amount of other articles and essays.

²⁶ R.M. Hare, Richard Brandt, William Frankena, John Rawls, Alasdair MacIntyre, Peter Singer, and Brad Hooker are just a few of the important theorists who believe Sidgwick to be a utilitarian, although

in terms of chronological order of influential works in this area of moral philosophy, and given Sidgwick's self-confessed alignment with Bentham on the definition of utilitarianism, his utilitarian status is entirely understandable²⁷. Epistemologically speaking, Sidgwick was an intuitionist - and this intuitionism too is thought to be utilitarian in nature²⁸. However, a cursory reading of the prefaces to the first two editions of *ME*, and of *ME*'s introduction itself, reveals something that should give immediate pause to the trend of simply labelling Sidgwick a utilitarian. This is Sidgwick's expressly made point that he is beginning his investigations in *ME* *not* from one position in particular that he will then defend, but precisely from a position of disinterested impartiality²⁹.

some of them do add caveats to that definition. An outline of utilitarian and other readings of Sidgwick appears in brief below, and in full in the next chapter.

²⁷ Sidgwick's moral theory can also generally be described as a dictate to aim at the general happiness, and on this basis too it is entirely reasonable that Sidgwick is interpreted to be a utilitarian. I do not specifically dispute here that he is a utilitarian of sorts. I do, however, argue that he should not be exclusively classed as a utilitarian, and that his utilitarianism is only established on a very specific and *deontological* basis. I make the outline case for this below – and in the thesis overall.

²⁸ See David Phillips, *Sidgwickian Ethics* (Oxford: Oxford University Press: 2011), Thomas Hurka, *British Ethical Theorists from Sidgwick to Ewing* (Oxford: Oxford University Press: 2014); Antony Skelton, Henry 'Sidgwick's Moral Epistemology', in, *Journal of the History of Philosophy*, Vol.48, No.4 (2010) 491-519 and 'Sidgwick's Philosophical Intuitions', in, *Ethics and Politics*, Vol.10, No.2 (2008) 185-209; D.D. Raphael, 'Sidgwick on Intuitionism', in, *The Monist*, Vol.58, No.3 (1974) 405-419; Robert Audi, *The Good in the Right* (Princeton: Princeton University Press: 2004).

Intuitionism, which generally holds that that moral propositions are self-evident and capable of being apprehended without reference to a process of reasoning, was among the dominant forms of moral epistemology from the 1700s (to the 1930s). But it is probably Sidgwick who entrenched it as a 'method of ethics', listing it as he did with the two other main methods, hedonistic utilitarianism and rational egoism. For Sidgwick, intuitionism is integral to how we come to know the moral principles that make up the 'morality of common sense', but Sidgwick's use of intuitionism is wide, varied, and sometimes inconsistent. Sometimes he argues that moral truths *can* be directly intuited, at other times he is mistrustful of such intuitions, and argues that they must be scientifically examined, and systemised. Intuitionism also appears in his accounts of all three ethical methods (common sense morality, utilitarianism, and rational egoism) and provides the means through which Sidgwick ultimately takes the step of arguing for congruence between at least two of those methods, as I will be arguing here.

Further to this point that Sidgwick deduced more than one moral principle from intuitionism, Phillip Stratton-Lake points out that ethical pluralism, in which there is an 'irreducible plurality of basic moral principles' with no principle taking precedence over another, is often considered to be a feature of intuitionism - but because not all intuitionists are pluralists he discounts it from his formal definition (Stratton-Lake, *Ethical Intuitionism: Re-Evaluations*, (Oxford: Oxford University Press: 2002). The two intuitionists who he references for evidence of this non-pluralism are G.E. Moore, and Sidgwick. It is not my primary aim in this thesis to directly expound the details of Sidgwick's intuitionism, but I do disagree with Stratton-Lake on this point. Sidgwick's personal account of the development of his ethical view (see below) very clearly shows Sidgwick accepting *two* fundamental intuitions – one deontological and one utilitarian – which were to become the central pillars of his moral theory. On this basis I also disagree with Phillips' and Hurka's interpretations of Sidgwick's intuitionism as being purely utilitarian. See below, and the literature review.

²⁹ See Sidgwick's claim in the first preface that *ME* is 'an examination, at once expository and critical, of the different methods of obtaining reasoned convictions as to what ought to be done' (*ME*: v), and the later passage, "I have wished to put aside temporarily the urgent need wShich we all feel of finding and adopting the true method of determining what we ought to do; and to consider simply what

This impartiality on Sidgwick's part as regards an ethical system is itself important to the claim I will be making throughout the thesis that Sidgwick should not simply be considered a utilitarian, as it is this that I believe leaves room for Sidgwick to present not just one moral theory that he was aiming all along to endorse, but one that actually seemed to develop naturally from his investigations into the various methods. But it is also important for emphasising Sidgwick's real motivation for writing *ME*, which accords directly with my starting premise of moral ambivalence. Sidgwick began *ME* from his observation that there is a plurality of moral principles employed by individuals, and these principles seem to be relevant either at different times or at once, and that they often conflict, both for individuals and between them³⁰. There is, Sidgwick remarks, a 'diversity of methods, applied in ordinary thought' (I:I:6). Sidgwick mentions this existence of varying and conflicting moral principles repeatedly throughout the introduction, with phrases such as 'the varieties of human conduct' (I:I:2), 'the diversity of answers which we find manifestly declared' (I:I:6), '...the unphilosophic man is apt to hold different principles at once, and to apply different methods in more or less confused combination' (I:I:6), 'different views of the ultimate reasonableness of conduct' (I:I:6), 'the different practical principles' (I:I:6), 'men commonly seem to guide themselves by a mixture of different methods' (I:I:12), and, perhaps most significantly, 'the original vagueness and ambiguity which lurks in the fundamental notions of our common practical reasonings' (I:I:13)³¹. All of these principles are, Sidgwick says, at one time or another 'satisfactory to the common sense of mankind', but only, he notes, 'as long as they have the field to themselves' (I:I:14). He summarises these observations on the last page of the introduction with a striking series of questions, clearly intended to invoke

conclusions will be rationally reached if we start with certain ethical premises" (*ME*: vi). He also claims at various points throughout the introduction (I:I:12, 13, 14), that his only aim is to expound as clearly as possible 'the different methods of ethics', that he found "implicit in our common sense reasoning". See also the preface to the second edition, where Sidgwick reasserts the fact, in the face of his critics (presumably F.H. Bradley) that he is *not* writing 'as an assailant of two of the methods...and a defender of the third' (*ME*: x). The significance of this originally intended neutrality – and the move away from it Sidgwick made in order to give his own moral theory, is discussed in chapter 4.

³⁰ During the period in which *The Methods* appeared, philosophy – both academic and popular – was highly confused. The outline as to why has already been given above in the discussion of the origins of the divide in moral philosophy: Kant's theory of absolute moral principles and duties had clashed dramatically with Bentham and Mill's positing of outcome of action being the only measure of moral value. Within this existing melee, other related issues – such as hedonism itself, egoism, rational egoism, intuitionism and virtue - were also being debated. Sidgwick wrote in direct response to this academic confusion, as well as the confusion in lived experience, and it is for this careful clarification and systemisation that *ME* was held in such esteem at its publication, and for a long time after.

³¹ See also I:VIII:100-102 for the idea that all of these approaches are included in our 'reasonings'.

introspection on the part of the reader: 'When I am asked, "Do you not consider it ultimately reasonable to seek pleasure and avoid pain?", "Do you not intuitively pronounce some actions to be right and others wrong?", "Do you not acknowledge the general happiness to be a paramount end?" "I answer 'yes'", he says, "to all these questions. My difficulty begins when I have to choose between them..." (I:I:14)³².

There is then one primary *raison d'être* for Sidgwick's writing of *ME*, and it is not his desire to defend utilitarianism. It is, quite simply, the situation common to the experience of all people - of moral ambivalence³³.

0.5.a. Sidgwick's Argument for Dual Reasonableness, and the Hope for Synthesis

Firstly then, Sidgwick is relevant to my arguments about moral ambivalence because he makes the very same points about there being more than one moral principle that seems to have a degree of truth, and yet these principles often conflict with or even directly contradict each other. But Sidgwick is relevant to the arguments made above in an even more significant way. This is that he argues the cause of the confusion to lie in the fact that these diverse principles all seem to have a claim to being simultaneously *reasonable*. He begins the introduction by defining 'a Method of Ethics' as *any* rational procedure by which we determine what individual human

³² He gives this answer from his own point of view, but it is clear that he believes himself, even as a philosopher, to represent all people on this point.

³³ Sidgwick's concern was with how philosophy was to go on in a climate in which definitions and theories of morality, the discussions over its dictates and nature, the apparent rightness or wrongness of all of them, and the experience of the ordinary human being were so utterly disordered. Appearing amongst the midst of the voices, theories, viewpoints and theorists all clamouring for prominence at this time, a visual impression that we might have from the first few pages of *The Methods* is of Sidgwick wading in to the fray, holding his hands up for everyone to just be quiet for a moment. For the task he had set himself of untangling this confusion of methods and principles, Sidgwick was ideal. As an academic and as a man, he tended to labour under the permanent conviction that there would always be another viewpoint, or another opinion, that might need to be considered. As his brother-in-law Arthur Balfour said: "Of all the men I have known, he was the readiest to consider every controversy and every controversialist on their merits. He never claimed authority; he never sought to impose his views; he never argued for victory; he never evaded an issue." (Sidgwick and Sidgwick, 1906: p. 311) See also the mission statement of The Sidgwick Society, set up in 1983 by Marcus Singer, which states itself to have been founded out of admiration for the character of Henry Sidgwick and the way he practiced philosophical inquiry: Careful, cautious contemplation, considering all points and every possibility, along with his resolute aim at truth and clarity, and clarified understanding" (in, Marcus Singer's introduction to Sidgwick's *Essays on Ethics and Methods* (Marcus Singer, (ed.), (Oxford: Oxford University Press: 2000), p. x.

beings ‘ought’ – or what it is ‘right’ for them – to do’ (I:I:1), and the language of ‘reasonable’ appears with noticeable frequency throughout the chapter (for example: “when a man asks ‘why he should do’ anything, he commonly assumes in himself a determination to pursue whatever conduct may be shown...to be reasonable’ (I:I:5), ‘the common practical reasoning of men generally’ (I:I:6), ‘rational ultimate ends’ (I:I:9), ‘our common moral reasoning’ (I:I:14))³⁴.

On closer inspection however, Sidgwick points out, the methods themselves are actually quite different. There is the Intuitional view, which Sidgwick defines as the position that holds moral rightness to lie in unconditional duties which are binding without qualification and without regard to ulterior consequences (I:I:3,8; see also I:VIII:98, 100); the view that rules are valid so far as their observance is conducive to general happiness (I:I:8); and the view that the proper end of individual action is private happiness (I:I:10). These three views essentially represent the three main ‘methods’ of ethics, that between them cause the confusion Sidgwick observes. Quite clearly, the first two viewpoints represent the deontological and utilitarian positions respectively, that were outlined above³⁵. But despite the differences, these positions still, for Sidgwick, share that common denominator of appearing to be reasonable. This is absolutely crucial to the arguments that Sidgwick will go on to develop throughout *ME* – for it is within what actions can be called *reasonable* that Sidgwick gives his definition of *rightness*. Further to this, and even more significantly, when Sidgwick first specifically discusses what is meant by ‘right’ he does not immediately give a traditionally consequentialist-utilitarian definition of ‘fit to an ulterior end’ (I:III:26). In fact, he initially negatively defines rightness in the distinctly un-utilitarian terms as *not* simply fit to an ulterior end. For Sidgwick, rightness

³⁴ As Sidgwick himself would admit much later, the question of conflicting reasons – of different principles that seemed nonetheless to be equally right - never left his mind as a philosopher. See also J.B. Schneewind for his elucidation of Sidgwick’s starting point that ‘we have a unique, irreducible concept of ‘being a reason for’, as it applies to desire and action’ (Schneewind, *Sidgwick’s Ethics and Victorian Moral Philosophy* (Oxford: Oxford University Press: 1977) 303-4.

³⁵ The last method is egoism, or rational egoism, to give it the full name Sidgwick assigned to it. Egoism was to prove the greatest challenge to Sidgwick’s notion that moral principles could possibly be assimilated. For Sidgwick, what it is ultimately rational to do appeared to apply just as much to individuals aiming at their own happiness, as it does to aiming at the general happiness. This problem became known as the Dualism of Practical Reason (hereafter DoPR) and it infamously dogged Sidgwick for most of the rest of his life. See below for a brief discussion of the DoPR, and why it is legitimate to ignore it for the purposes of my argument in this thesis. Also see below for my justification of the link between Sidgwick’s definition of ‘Intuitionism’ in this sense, and the deontological position.

actually lies in what can be proven to be ultimately *rational* – in either certain moral duties thought to be unconditional and thus intrinsically right in themselves, or in certain ends that can similarly be proven to be ultimately rational. But both of these cases – certain actions or certain ends – share that non-consequentialist basis of rightness. Here, Sidgwick’s definition of right as being that which is reasonable actually accords very closely with Kant’s strictly deontological argument for the relationship between rightness and Reason³⁶. And Sidgwick agrees further with Kant that the recognition of something as ultimately reasonable will, for rational creatures, naturally bring about the impulse to pursue it; that is, it incurs an ‘ought’³⁷. The pursuit of what is recognised to be ultimately rational thus becomes, for Sidgwick, an unconditional categorical imperative – and in this takes on the quality of rightness as Sidgwick understands it. Of course Sidgwick differs from Kant in ascribing the categorical imperative not just to certain moral actions (duties), but to the adoption of certain ends, also considered to be ultimately rational, such as the Universal Good, or ‘Happiness’³⁸. But in both cases, he says, we do not ‘get rid of the dictate of reason’ (I:III:36). That is, Sidgwick attributes reasonableness, and ultimately

³⁶ I am following Sidgwick’s lead in capitalising ‘Reason’ here (and throughout the thesis), in order to make clear that this term refers to the faculty that is essential for determining moral conduct, as Sidgwick describes it.

Sidgwick justifies his use of ‘Reason’, and its relationship to ‘rightness’ that I am also assuming in this thesis, at I:III:34. He states here that does not mean to prejudge that all moral judgements are derived from a process of reasoning, or universal principles, admitting that we often deal with moral situations on a more case by case basis. However, according to Sidgwick this implies the ‘moral sense’, which in turn suggests ‘feelings’ - and feelings cannot truly be in error, objectively speaking, however much they may vary from person to person. Sidgwick therefore elects to use ‘Reason’ to ‘denote the faculty of moral cognition’ (I:III:34). This faculty tends to generate general rules, to which individuals appeal when our moral judgements are confused, unclear, or conflicted. ‘Reason’ is thus the mechanism through which we discern what it is ‘right’ to do (see I:III:35).

³⁷ For the full discussion of the relationship between Sidgwick’s work and Kant’s, see chapter 2.

³⁸ There is, however, is evidence to suggest that actually this might not be so ‘obviously’ un-Kantian as it might first appear. Kant’s definition of ‘highest good’ is also ‘happiness’, but *only*, he stipulates, as conceived by rational agents (*Critique of Pure Reason*, 1781 6:5-7). As will be seen in chapter 2, Sidgwick is extremely specific about the establishment of Happiness as a rational end, and about how it is only this prior establishment of ultimately rational ends that can make the means to those ends in any way ‘right’. Where they do differ is that the pursuit of Happiness is not for Kant an additional duty, but simply the sum of all other duties (*Critique of Pure Reason*, 1781: 6:5). The emphasis on unconditionality, and the relationship between that and the ‘ought’ of moral duty, in Sidgwick’s definition of rightness forms a significant part of my argument made in chapter 2 that Sidgwick could be called a deontologist to some extent. There has also been some literature – the work, for example, of Mariko Nakano-Okuno (‘Sidgwick and Kant’: On the So-Called Discrepancies Between Utilitarian and Kantian Ethics’, in, P. Buculo, R.Crisp, and B. Schultz (eds.) *Henry Sidgwick: Happiness and Religion* (Catania: Catania University Press: 2007) 260-332) - arguing for greater similarities between Sidgwick and Kant than might commonly be assumed. This literature provides important support for my argument that Kantian ideas have a distinct influence in Sidgwick’s work, but the argument has not yet been made that Sidgwick could actually, on these grounds, technically be called a deontologist (or even a Kantian), or attention drawn to the specifically Kantian basis of his utilitarianism.

rightness to *both* deontological and utilitarian considerations³⁹. This strong emphasis on Reason, and the subsequent dual ‘rightness’, will be at the heart of Sidgwick’s own moral theory.

When outlining in I:I the approach he has decided to take to the methods of ethics, Sidgwick states that ‘we cannot regard as valid reasonings which conflict’, and that ‘two conflicting rules of action cannot both be reasonable’. Sidgwick is right that this would not seem to conform to the definition of ‘reasoning’. He also postulates that when two methods conflict, one or the other must be modified or rejected (I:I:6). But Sidgwick does also consider another possibility. He acknowledges that we tend to feel some need to harmonise these alternatives between which – so far as they cannot be reconciled – the human mind seems...forced to choose’, and that mentally at least, the human mind itself will attempt a *synthesis* (I:I:12) between these practical maxims. The result is usually a ‘confused blending, or a forced and premature reconciliation’, and writers who have attempted such systems have usually ‘proceeded...without adequate analysis’⁴⁰. Although Sidgwick denies that he will be attempting an actual solution to the difficulties and conflicts between the methods, he does express two points of hope pertaining to this idea. The first is that he hopes to ‘afford aid towards the construction of such a system’; the second is that ‘we cannot but hope that all methods may ultimately coincide’ (I:I:14). And these passages appear alongside Sidgwick’s language of reconciliation, and synthesis. Where Sidgwick did not see grounds for reconciliation, this was between egoism and the other two methods. As for the deontological and the utilitarian positions however, Sidgwick will not, in the end, divide them on the basis of their shared appeal to Reason - but actually *unite* them on that very basis, in his own moral theory. It is my view therefore that Sidgwick’s hope for these apparently valid and yet conflicting reasonings to

³⁹ In Sidgwick’s final work before his death in 1900, *Practical Ethics*, he gives a particularly illuminating statement regarding his view of Reason, and its primary role in establishing moral obligation: “Moralists of all schools have acknowledged....that broad agreement [reason] in the details of morality...Well, my view is that we ought to start with this broad agreement as to the dictates of duty...the first principles on which duty may be constructed as a rational system, to make this general agreement somewhat more explicit and clear than it is in ordinary thought...” (Sidgwick, *Practical Ethics*, pp. 6-7).

⁴⁰ Nor, Sidgwick points out, have attempts to disguise the differences worked; this tends to result in nothing more but heightened controversy later on. As Sidgwick succinctly summarises, ‘Thus we get on the one hand vague and hazy reconciliation, on the other loose and random exaggeration of discrepancies; and neither process is effective to dispel the original vagueness and ambiguity which lurks in the fundamental notions of our common practical reasonings’ (I:I:13)

ultimately coincide is actually fulfilled, and that it is embodied in his theory of Rational Benevolence.

0.5.b. 'The truth lies between these two conclusions': Sidgwick's Moral Theory of Rational Benevolence

It is clear then from even I:III that Sidgwick has a very definite understanding of what is *right*. If we then follow the progression of *ME*, a process other than an investigation into what people *do* do and think begins to emerge – a process in which Sidgwick begins to argue for what people *ought* to do. That is, *ME* does – despite Sidgwick's original claim to impartiality – actually contain a normative moral theory⁴¹. This is brought about by Sidgwick perceiving there to be that fundamental connection between those deontological principles and utilitarian principles, which is established by his argument that both types of principle are ultimately rational. I have already stated that this fundamental connection is at the centre of Sidgwick's moral theory, but the construction of that theory – i.e. the process that leads to Rational Benevolence - is also important. In I:III it is seen that Sidgwick believes the moral cognition of Reason to generate general rules, or moral obligations. In I:VIII, Sidgwick then draws out what he considers to be the three forms of the Intuitionist method, the second of which – dogmatic Intuitionism – represents that collection of rules (also understood to represent the deontological position, and also called the morality of common sense)⁴². These rules are not, however, by themselves adequate to provide a full system of moral guidance. This is firstly because those

⁴¹ What is meant by 'moral theory' here is a set of statements or principles used to explain moral standards, and/or a system of guidance for implementing those standards. The explanatory capacity of a moral theory pertains mainly to meta-ethical theories, and the guidance capacity to normative theories. I project that Rational Benevolence qualifies under both of these types, although the aim of the thesis is mainly to develop the meta-ethical aspect, rather than the normative. The normative value of Rational Benevolence is, however, briefly explored in the thesis conclusion.

⁴² It is here that we see Sidgwick's direct connection between dogmatic Intuitionism, or common-sense morality, and his identification of the Intuitionist method as representing the deontological position. Sidgwick unambiguously states at I:VIII:98 that he takes the 'distinguishing characteristic of the Intuitionist method' to be that there are 'certain kinds of actions unconditionally prescribed without regard to ulterior consequences' (which obviously delineates this position from the utilitarian position). My argument that Intuitionism in *ME* represents a deontological (and often recognisably Kantian) position, and my subsequent point that *this* type of Intuitionism (represented by a capital 'I') is actually a crucial part of Sidgwick's own epistemic intuitionism (represented by a small 'i'), appear in chapter 2. The need to delineate these two uses of intuitionism in this way is explained below.

rules still only make up ‘an accidental aggregate of precepts, which stands in need some rational synthesis’, but – more seriously – it is because they are not actually self-evident. That is, they are not ultimately rational first principles⁴³. For this reason, they do not incur an ‘ought’, and are not ‘right’ in the absolute sense that Sidgwick means by that term.

Having drawn this rather gloomy conclusion, Sidgwick then states – in a passage that is highly revealing of the direction in which his theory is going to go - that in the search for ultimately valid and true moral principles ‘we shall find, I think, that the truth lies between...two conclusions. There are certain absolute practical principles, the truth of which...is manifest; but they are of too abstract a nature, and too universal in their scope, to enable us to ascertain by immediate application of them what we ought to do...; particular duties have still to be determined by some other method’ (III:XIII:379)⁴⁴. Next, in a chapter entitled ‘Philosophical Intuitionism’ (which was the type of epistemological intuitionism Sidgwick predicted would be necessary in the end), Sidgwick turns his attention to what *can* actually be called self-evident – to what can be called ‘right’ in the absolute sense. It is important to note here that he does not simply abandon the idea of Intuitionism’s self-evident principles. The first of Sidgwick’s own self-evident principles is an inverted form of Kant’s universalising maxim, which Sidgwick calls the principle of Justice. Being self-evident, ultimately rational, and therefore wholly self-sufficient for its own justification, Justice is, I argue, a non-consequentialist principle. Into the foundational framework of Justice, Sidgwick then adds the logical progression to the self-evident principle of Prudence, which holds that it is only rational for individuals to aim at the *whole* of their good, rather than just one part of it. Again, given that this is an ultimately self-evident principle, Prudence is also non-consequentially right without further reference to anything else. Finally, Sidgwick deduces that if our own individual good, at which it is rational to aim, is a sum of parts, then the general good, which is a sum of the parts of all individuals, is also therefore an end at which it is rational to aim (III:XIII:379-382). This ultimately rational general good thus incurs an ‘ought’, and becomes – as

⁴³ As will be seen in chapters 2 and 4, Sidgwick believes Kant’s original universalising maxim to be at fault here.

Sidgwick disputes the self-evidence of Intuitionism’s/common sense morality’s principles in Book III via an eleven-chapter investigation into those rules, and argues in the end that only Justice, Prudence, and Benevolence truly admit of being ultimately rational, and therefore self-evident.

⁴⁴ It is almost universally agreed that this ‘other method’ would be utilitarianism. I will be arguing against this view that utilitarianism is simply superimposed over common sense morality.

per Sidgwick's earlier definition of ought and right – a categorical imperative, represented by the maxim of Benevolence (also called Universal Benevolence)⁴⁵. Defined as 'each one is morally bound to regard the good of any other individual as much as his own', Benevolence is clearly a utilitarian principle. However, it is crucial to note that Sidgwick has established Benevolence on the basis of principles which can be seen to be predominantly non-consequentialist and thus *deontological* in nature - and that this is the only way that Benevolence can be fully rationally supported. That is, Sidgwick required these particular grounds in order to *derive* the utilitarian principle. Furthermore, as an ultimately rational end Benevolence itself is a non-consequentialist moral obligation, or duty, because – like Justice and Prudence - being ultimately rational it does not pertain to anything further⁴⁶. It is in this way that Sidgwick's moral theory 'lies between the two conclusions' of self-evident deontological principles, and the utilitarian principle. His theory includes at once a distinct deontological aspect, and a distinct utilitarian aspect, both of which share the validity given by ultimate rationality, and both of which are crucial to a full understanding of Universal Benevolence. More specifically then, this Benevolence – and this moral theory - is more appropriately represented by the term 'Rational Benevolence'.

0.5.c. Reconciliation in Sidgwick: Rational Benevolence as a Synthesis between Deontology and Utilitarianism

⁴⁵ Sidgwick capitalises the form of 'Benevolence' that is a self-evident principle, that emerges from his process of deducing truly rational principles that I have just outlined. For that reason, I will also capitalise Benevolence throughout the thesis. Sidgwick sometimes uses Benevolence interchangeably with 'Universal Good', and for the same reasons 'Universal Good' will also be capitalised. Otherwise, 'good' is not capitalised. This also follows Sidgwick's example. In terms of quite what Sidgwick means by this general good, he later defines it as 'Happiness' (III:XIV:391-407). Sidgwick is not always consistent about capitalising 'Happiness', but for the most part it seems as though he capitalises it when it is discussed as being the end of the maxim of Benevolence. Happiness will therefore usually be capitalised throughout the thesis, to represent the idea that it is a rational, ultimate end, as I will argue in chapter 3.

⁴⁶ As was outlined above, duty is defined here as the non-consequentialist position, in order to avoid the confusion that might arise between deontologically conceived duties, and utilitarian type duties. It is important to note that by 'duty', Sidgwick is referring to the non-consequentialist position. To say that 'duty' can apply just as much to utilitarian principles is actually not an entirely inaccurate representation of Sidgwick's final position on moral obligation. However, I do not think that this is the view of duty from which Sidgwick *begins* in *ME*, or the one that he uses to establish his moral theory in the first place. In fact, I will argue in the thesis that his theory of moral obligation is based on a notion of duty, or 'ought', that is unconditionally and categorically prescribed through Reason without regard to ulterior consequences. Chapter 2 will argue this in detail.

I argued in the first half of this introduction that the most effective and suitable way to address the problem of moral ambivalence would be to develop a theory that shows both absolute-principle-based (deontological) and outcome-based (utilitarian) moral principles to be reasonable, and that the most effective way to do *this* is to show that they are actually inherently connected, in a mutually informing synthesis. I have given the outline as to how Sidgwick includes rational deontological moral properties and rational utilitarian properties together in his moral theory of Rational Benevolence, but it must now also be theorised as to how these principles might be mutually informing, as per the definition of synthesis given above. A brief overview of the case for the synthesis is as follows⁴⁷.

Firstly, I argue that the Benevolence in Rational Benevolence rests on prior, fundamental, and non-consequentialist principles – represented by Justice and Prudence. It is the logical connection between these principles that then also logically yields the utilitarian principle of Universal Benevolence. In fact, Sidgwick states Rational Benevolence to be the rational basis – the *proof* – of utilitarianism, without which utilitarianism is lacking the crucial, unconditional ‘ought’ which makes it *right*, on Sidgwick’s understanding of that concept⁴⁸. This language of proof may appear to elevate the deontological properties above utilitarian ones, or, alternatively, create the impression that once utilitarianism has been derived from deontological properties it can then be held up as the ultimate moral theory, which is now ‘proven’. It is true that the utilitarian principle is derived from other principles – and it is true that those principles provide the rational basis that Sidgwick thought to be required for the utilitarian system. But it is *not* the case, I argue, that utilitarianism simply ‘trumps’ those principles. This is for two reasons. The first is that without those principles retaining their non-consequentialist properties, the utilitarian principle could not be established in the first place. It would be detrimental to utilitarianism therefore for that principle to try and simply supersede Justice and Prudence. Secondly, instead of the utilitarian principle of Universal Benevolence

⁴⁷ Obviously it is the remit of the main thesis to fully develop this argument, but identifying the basic stages of the argument creates in turn a series of stages through which it – and the thesis – will need to pass, in order to present a persuasive case for Rational Benevolence as a synthesis. These stages appear below, where the structure of the thesis is given.

⁴⁸ Sidgwick considers this version of proof to be that which even Mill’s doctrine still lacked, despite Mill’s attempt to identify it. See chapter 4.

overriding the other two principles, it actually *contextualises* them. As was seen in Sidgwick's claim that 'the truth lies between these two conclusions', there *are* certain absolute principles, the truth of which is manifest - but they require systemisation by another, more specific principle. It is in this way that the mutually informing synthesis is truly created. Where the deontological properties of Justice and Prudence are indispensable to the establishment of utilitarianism, utilitarianism gives those two principles full *meaning*, as part of a comprehensive moral theory. Deontological principles provide the proof of utilitarian ones, and those utilitarian ones fully support and validate the deontological principles. I argue that this result is a circularly interdependent relationship in which neither moral property is complete without reference to the other – and in which morality is neither wholly deontological (which might be the interpretation if we only see that utilitarianism is *derived* from a deontological basis), nor wholly utilitarian (which might be the interpretation if we only see that that utilitarianism then confirms the role of the deontological principles). The two types of moral property are not, however, simply the *same*. As well as each being informed by the other, they both also retain their own respective deontological or utilitarian identities; without this degree of distinctiveness, neither property can provide the support needed for the other, and both collapse. Thus, deontological and utilitarian properties are fully reconciled in an inextricable way, within this mutually informing and dynamic relationship. Both have value, both are ultimately reasonable, both look to the other for validation – and both are seen to be vital to the establishment of Sidgwick's theory of Rational Benevolence.

0.6. Sidgwick as a Deontological/Utilitarian Intuitionist, and the Personal Document

There have now emerged two major claims that are central to my overall argument that Sidgwick developed a synthesis between deontological and utilitarian properties of morality, both of which must be addressed in order to develop that argument in full. The first is the as yet somewhat clandestine assertion that Sidgwick was, to a recognisable extent, a deontologist. I recognise that this is probably the most outlandish claim of the thesis. But it is, in my view, more than defensible. Sidgwick's moral theory of Rational Benevolence relies on a principle of non-consequentialist moral obligation, and this is the understanding of moral obligation that is traditionally

found in the deontological school. It is true that his theory of what is right being that which is rational allows the notion of 'ought' to transcend the traditional deontological/utilitarian boundaries of Sidgwick's day, but, as I will argue, Sidgwick does not then just attribute this 'ought' to the utilitarian principle, and then discard its deontological foundations. He maintains an independent notion of non-consequentialist duty that is in its own right essential to Rational Benevolence – without *that* particular understanding of 'ought', utilitarianism cannot be established in the first place.

The second claim is that Sidgwick's utilitarianism is not in every sense as traditional or as classical as it is often assumed to be, and that this is crucially what allows, and even *requires*, that his utilitarianism be combined with deontological principles. Related to, and uniting, both of these points is the issue of Sidgwick's intuitionism, which it is conducive here to use as a framework for making the wider argument about Sidgwick's deontological/utilitarian moral epistemology.

There has been, over the years, some confusion over Sidgwick's use of intuitionistic terminology, given that on one hand he uses it to describe a particular method of ethics, and on the other uses intuitionism himself as an epistemic device. So prevalent was the confusion in fact, that in *Five Types of Ethical Theory*, C.D. Broad introduced the term 'deontology' in direct response to Sidgwick, in an attempt to distinguish between what Sidgwick means by 'Intuitionism' as a method of ethics and 'intuitionism' as an 'epistemic principle of classification'⁴⁹. Given my claim that an unhelpful and inaccurate divide has been maintained between deontological and utilitarian principles of morality, and the *central* claim that Sidgwick's moral theory can provide some kind of solution to this, it is interesting – and even ironic – that those very terms, and that very distinction, should have arisen (with Broad) as a direct result of Sidgwick's work. But I actually agree with Broad's classification. For by 'Intuitionism' with a capital 'I' (as I propose to distinguish it) Sidgwick does indeed mean the traditional, non-consequentialist deontological view⁵⁰. But then there is the issue of his own epistemic intuitionism (written with a small 'i'). David Phillips, whose 2011 book *Sidgwickian Ethics* is among today's most incisive accounts of Sidgwick's

⁴⁹ Broad, *Five Types*, 162-163.

⁵⁰ I argue this point in full in chapter 2, but here it suffices to point out that Sidgwick describes Intuitionism as the view that 'conduct is held to be right when conformed to certain precepts or principles of Duty, intuitively known to be unconditionally binding (I:I:3), and 'without qualification and without regard to ulterior consequences' (I:I:8)

own methods in *ME*, concisely refers to Sidgwick's particular use of intuitionism as 'intuitionistic utilitarianism'⁵¹. Phillips is not alone in equating Sidgwick's intuitionism exclusively with a utilitarian position; Antony Skelton also does this, as does Thomas Hurka. I will not in any way be disputing that Sidgwick is an intuitionist - Sidgwick is probably more clear that he is an intuitionist than he is anything else, and his moral theory is clearly intuitional, epistemologically speaking. But, as I do agree that intuitionism is indeed Sidgwick's main moral epistemological method, my claim that there is a significant deontological element in Sidgwick's moral theory would mean claiming that his intuitionism is also partly deontological – and this is precisely what I claim. That is, I dispute that his intuitionism is *only* utilitarian. Sidgwick's *own* epistemic intuitionism, that informs his theory of moral obligation, is, on my view, actually itself derived from the deontological, Kantian terms with which he defines Intuitionism as the deontological-based method of ethics. For that reason, his wider epistemic 'intuitionism' sometimes closely accords with his 'Intuitionism', understood more narrowly as representing the deontological method - and this is where I believe the confusion to have arisen⁵².

I do not believe that this view (that Sidgwick's Intuitionism is also Sidgwick's intuitionism) has been fully considered. Phillips has argued, rightly, that that confusion is due to Sidgwick's using the term to represent two interpretations of 'intuitionism' (as a representation of a particular method of ethics, and the epistemic). Sidgwick does do this in places, but according to Phillips there is no other connection between the two uses. But it is more than possible to argue from *ME*, as I will do in chapter 2, for Intuitionism as representing a traditionally deontological, and even Kantian, position; it is also possible in *ME* to argue for that relationship between

⁵¹ Phillips, *Sidgwickian Ethics*, p. 11

⁵² I do agree with Broad's point that there needs to be a certain amount of delineation between, as David Phillips would say, the 'moral theoretic and the epistemic' usages of intuitionist terminology by Sidgwick in the later stages of *ME*, and it is for this reason that I propose using the upper-case/lower-case 'I's, as Sidgwick himself does. It is clear at times that Sidgwick believes all three main methods in *ME* (Utilitarianism, Egoism and Intuitionism) to be informed by intuition, whereas his term 'Intuitionism' is intended to only refer to one particular method (the non-consequentialist deontological), and this has confused matters. As Sidgwick himself admits in 'Professor Calderwood On Intuitionism in Morals' (*Mind*, Vol.1, No.4 (1876) 563-566) he found it very difficult to define intuitionism in an adequately 'useful' way; simply stating that 'moral principles are intuitively known' did not, for Sidgwick, distinguish Intuitionism clearly enough from Utilitarianism (p.564). However, it is my view that some of the confusion has arisen due to the fact that Sidgwick's 'intuitionism' actually shares a fundamental principle with his 'Intuitionism'. I outline this point here, and make the case fully in chapter 2.

‘Intuitionism’ and ‘intuitionism’, by claiming that Sidgwick’s epistemic intuitionism includes both deontological *and* utilitarian qualities.

Admittedly, Sidgwick is not explicit in *ME* that it is Kant’s doctrine to which he is referring when he defines Intuitionism in that way - but he is explicit about Kant’s role in his work in another document⁵³. This supplementary document is absolutely crucial to this understanding of the deontological/utilitarian character of Sidgwick’s intuitionism. Entitled ‘Professor Sidgwick’s Account of the Development of his Ethical View’, (hereafter referred to as the ‘personal document’, or ‘PD’), this autobiographical fragment was found among the manuscripts that Sidgwick had intended for publication along with the sixth edition of *ME* in 1901. Sidgwick died shortly before that publication, and as he had been for some time before his death unable to continue the revision of *ME* that he had started, he requested that pages 277 to the end were sent to the press unchanged. This was overseen by his devotedly loyal student Emily E. Constance Jones, who also made the decision to include the fragment in the preface to the sixth edition. Jones describes the piece as being both ethically and historically ‘of very exceptional interest’, and she is absolutely right. In this enlightening little narrative, Sidgwick takes the reader through the process he underwent while searching for an ethical position to which he could truly adhere. The first ‘stop’ in this journey is at Kant’s universalising maxim, which provided for Sidgwick a fundamental intuition that he felt was required *before* even the utilitarian principle could be accepted. Later in the document, Sidgwick explains how the Kantian maxim combines with a utilitarian intuition, to create Sidgwick’s eventual position – an intuitionism that is comprised of both a deontological/Kantian principle, and a utilitarian one. He describes this position in these words: ‘I was an Intuitionist again, but on a utilitarian basis’⁵⁴.

⁵³ In fact, as Alan Donagan points out, the Intuitionist method (associated with dogmatic intuitionism) which Sidgwick criticises in Book III, is actually likely to be Whelwell’s version of intuitionism (Alan Donagan, ‘Sidgwick and Whelwellian Intuitionism; Some Enigmas’, in, *Canadian Journal of Philosophy*, Vol.7, No.3 (1977) 447-465). I argue however that Sidgwick *does* retain the ‘unconditionally prescribed without regard for consequences’ aspect of Intuitionism in his own moral theory – and that in doing so he actually uses Kant. See chapter 2.

⁵⁴ In the PD we do see that Sidgwick associates the Intuitionism that I will be arguing accords with a Kantian version of deontology with writers other than Kant – but when he makes the statement that he is ‘an Intuitionist again...’, he continues to capitalise the word, and we know from the preceding material in the document that the Kantian principle has provided him with that intuitionism. This forms part of the grounds on which I will argue that Sidgwick’s ‘Intuitionism’ with a capital ‘I’, is a crucial part of his epistemic ‘intuitionism’ with a small ‘i’.

Shortly after this statement, Sidgwick declares ‘the opposition between Utilitarianism and Intuitionism was due to a misunderstanding.....I could find no real opposition between Intuitionism and Utilitarianism’ (*ME*: xx)⁵⁵. And *this*, he says, was the state of mind in which he published his book – from the point of view that the common conflict between Utilitarianism and the deontological position could not only be discarded, but overcome in such a way that they are actually reconciled.

The PD is, I believe, of the utmost importance for a full understanding of Sidgwick as a moral philosopher. It clearly demonstrates the fact that he attributes value to both deontological and utilitarian principles of morality, and that this informs his moral theory⁵⁶. It supports the view that Sidgwick’s moral theory – and his intuitionism - could only proceed on a fundamental intuition that was Kantian in origin, and thus wholly supports my argument that there is a fundamental deontological element to Sidgwick’s theory of Rational Benevolence. It clearly shows the link between the non-consequentialism of the Intuitional position in Sidgwick, and the non-consequentialism in Sidgwick’s own epistemic intuitionism – a link that is created by the Kantian maxim, and that gives the grounds for my claim that part of Sidgwick’s ‘Intuitionism’ also appears in his ‘intuitionism’. It also provides the most irrefutable evidence that Sidgwick saw a fundamental relationship between deontological and utilitarian principles, and that on his view to divide them is a mistake⁵⁷. Lastly, it gives good grounds on which to argue that Sidgwick should not simply be read as a classical utilitarian, who is confined to that field. Where he states in the personal document that he is a utilitarian, this is only on a very specific basis, and this is a deontological one⁵⁸. I will, at various points throughout this thesis,

⁵⁵ Sidgwick continues on to outline the reciprocal relationship between Utilitarianism and the Morality of Common Sense here, and these passages correspond directly with passages in *ME* itself, but I reserve discussion of this point for the main argument in chapter 4.

⁵⁶ The original document still exists today, in the special collection at The Wren Library at Sidgwick’s own Trinity College, Cambridge, complete with Sidgwick’s margin notes, written in his tiny, decorative handwriting. See also his correspondence with E.E.C.J herself.

⁵⁷ The claim that Sidgwick’s own moral theory is deontological to a degree obviously challenges Broad’s efforts to distinguish between Sidgwick’s deontological and teleological based methods. As noted above, Broad’s distinction contributed considerably to the divide between the two methods of ethics, and his application of it to *ME* naturally brought Sidgwick into the fold of theorists who had perceived, developed, and maintained this divide. The claim that Sidgwick’s own epistemic intuitionism retains a degree of deontology provides grounds for the thesis’ position that Sidgwick’s work – which is also utilitarian - is *not* guilty of perpetuating that divide, and that it actually goes a significant way towards reconciling it.

⁵⁸ So specific, in fact, is this basis, that I suggest that he should not be referred to as a ‘utilitarian’, given the suggestion of that label that this position must exclude him also being in part a deontologist. However, Sidgwick himself did not expressly dispute his utilitarian label, and it would therefore be presumptuous to do so here. I do maintain however that Rational Benevolence is *not* simply

theorise that he could also perhaps be described as a deontologist on a utilitarian basis. On my account of how Sidgwick developed this position into his moral theory of Rational Benevolence, either understanding is, I believe, credible.

0.7. Sidgwick's Place in History

Despite the importance of the PD as a vital supplement to the arguments of *ME*, there is relatively little interest in the document. It is not entirely ignored; most Sidgwick scholars refer to it at some point (Bart Schultz, David Phillips, Peter Singer, and Antony Skelton, for example). But I maintain that its real significance for improving understanding of Sidgwick's argument for the relationship between deontological and utilitarian properties of morality has not been fully realised. As a result, that argument as it appears in *ME* has also largely been overlooked in the 20th century— and other aspects of his work have been emphasised instead. Here, I first suggest the possible reasons for this. This leads on to the next part of the discussion, which takes account of the ways in which Sidgwick's work *has* been recognised or utilised, and thus begins the process of framing the place for my arguments made in this thesis within the wider fields of both Sidgwick studies and moral philosophy.

0.7.a Sidgwick at the Turn of the Twentieth Century

Among the first factors to unfavourably affect *ME*'s long-term reputation was the period in which Sidgwick was working. British Idealism had been on the rise from the mid 1800s, and *ME* struggled to defend its hedonistic emphasis against that tide. T.H. Green and F.H. Bradley were particularly influential in this movement, and where Green sympathised with some of Sidgwick's arguments, Bradley's criticism of *ME* was fierce and protracted⁵⁹. Even more significant was the changing socio-

utilitarianism, and that for this reason it should be referred to as Rational Benevolence, and not as 'Sidgwick's utilitarianism'.

⁵⁹ Some of Sidgwick's contemporaries – T.H. Green included - did actually recognise the nuances of the position that Sidgwick presented in *ME*. For F.H. Hayward, the point was not that Sidgwick departed from utilitarianism towards intuitionism, but that 'it was extremely difficult to classify him at

economic climate of this time. The turn of twentieth century, and the fifty years that followed it, was to be a time of great revision for both deontology and utilitarianism, and these modifications were often made in light of the emerging social and political issues. Sidgwick was, as Ross Harrison enduringly describes him, a ‘philosopher’s philosopher’, and *ME* is thought to reflect this via the dry and dull style in which Sidgwick apparently also lectured⁶⁰. *The Methods of Ethics* thus seems to stand on a certain horizon between the 19th and the 20th centuries – the classical ‘fathers’ theories behind it, and vast social, political and philosophical change spread out before it. As soon as those changes began, which they did immediately and in earnest, the deeply academic and theoretical *Methods* rapidly lost relevance, and faded into the distance.

The Dualism of Practical Reason had also, of course, caused a problem – to Sidgwick at least. Although *ME* is certainly not remembered just for this feature, for

all’ (Hayward, ‘Mr Hayward’s Evaluation of Professor Sidgwick’s Ethics: A Reply’, in, *Mind*, Vol.11, No.3 (1901) 360-365). (See also James Seth (‘The Ethical System of Henry Sidgwick’, in, *Mind*, Vol.10, No.38 (1901) 172-187, who believed Sidgwick to be neither utilitarian, nor intuitionist). Hayward also prophesied about *ME* that ‘to neglect it is a philosophical disaster’, and argued that a reader of *The Methods* must approach it not expecting an answer, as such. ‘The lesson that Sidgwick has to teach us’, Hayward said, ‘is the difficult lesson of openness of mind, of freedom from dogmatism’ (Hayward, ‘Evaluation’, p. 2). See also Hayward’s ‘The True Significance of Sidgwick’s Ethics’, in *International Journal of Ethics*, Vol.11, No.2 (1901) 175-187. These readings of Sidgwick and their long-term effects for *ME* are explored further in the literature review.

⁶⁰ Ross Harrison (ed.), *Henry Sidgwick* (Oxford: Oxford University Press: 2001) p.4. Despite the esteem in which many readers held *ME*, this rather less complimentary impression is also a common one. Not everyone was enthused by Sidgwick’s rigorously methodical approach; even Hayward admitted that *ME* had its dry and challenging reputation for a reason. Sidgwick’s style was apparently especially difficult when it came to being taught by him. Many students of Sidgwick – among them F.H. Bradley, G.E. Moore and Bertrand Russell, who represented almost all of Sidgwick’s critics and who would also go on to write their own moral theories – reported that Sidgwick told precisely one joke per lecture. As soon as it had been and gone, many of the students would, if they could, quietly leave (Russell, *My Philosophical Development* (London: Allen and Unwin: 1959) p. 38. See also Ross Harrison, ‘Cambridge Philosophers IV: Henry Sidgwick’, in, *Philosophy*, Vol.71, No.277 (1996) 423-438.) But it is unfortunate that both Sidgwick’s lectures and *ME* have been remembered in this way. Not only is *ME* actually highly engaging in places, but in many other accounts from Sidgwick’s pupils a very different picture of him is created, both as a professor, and as a man. In conversation with his students, (which he preferred to lecturing) he was known to be witty, and gently humorous in any counter arguments. He was apparently sharp but sympathetic, and was loved for his candour – or, as his students termed it, his ‘Sidgwickedness’ (Bart Schultz, *Henry Sidgwick: Eye of the Universe*, (Cambridge: Cambridge University Press: 2012) p. 68). In the Sidgwick archives at Newnham College in Cambridge, there is a wonderful anecdote from one of his female pupils, who was among the first to be reading there following Sidgwick and his wife Eleanor’s instrumental role in having women admitted to Cambridge University. She had received a paper back from Sidgwick, and was unable to read his handwriting. While everyone else in the quiet classroom continued to work, she approached Sidgwick, reading behind his desk, and asked if he could tell her what this word said. Sidgwick looked down at the word, looked up at her over the top of his spectacles, and said with a wry smile, but not quite loudly enough that everyone could hear “that little word is b – a – d -bad”.

Sidgwick the DoPR was a crisis of almost catastrophic proportions⁶¹. The very reason for this crisis was that Sidgwick believed himself to have failed to successfully assimilate all rational methods of ethics, with egoism and altruism still appearing fundamentally at odds. Sidgwick allowed this impression to prevail, both in his own mind and in his reputation, and as a result the ‘failure’ arguably overshadowed the success of *The Methods* – and obstructed any sense that he might have reconciled other methods⁶².

Finally, I believe that one of the main reasons for the overlooking of Sidgwick’s reconciliation is that today *ME*, now in its 7th edition, is rarely published with the prefaces included. This is to its great detriment, for there are several passages in the first two prefaces (especially in the second, where Sidgwick responds to critics of the first publication) that could have clarified Sidgwick’s position on some important points⁶³. In the absence of these brief but important qualifications, one idea in

⁶¹ To use Sidgwick’s most dramatic imagery, there is a ‘fundamental contradiction in our apparent intuitions of what is Reasonable in conduct’ such that ‘the ‘Cosmos’ of Duty is thus really reduced to a Chaos, and the prolonged effort of the human intellect to frame a perfect ideal of rational conduct is seen to have been foredoomed to inevitable failure’ (*ME*1:473). This passage was the original closing paragraph of the first edition of *ME*. He was later persuaded to change it – but the fact remains that originally, Sidgwick was prepared to conclude on *ME* that its first word was ‘ethics’ and its last was ‘failure’.

⁶² Sidgwick refers briefly to the problem of rational egoism in the PD, stating that ‘there was indeed a fundamental opposition between the individual’s interest and either morality, which I could not solve by any method I had yet found...’ (*ME*: xx). But it does not seem to have the same depressing effect on him here as it did in his personal correspondences on the matter. David Phillips actually argues that Sidgwick was *less* successful than he (and myself) thought in reconciling utilitarianism with dogmatic intuitionism, and *more* successful in reconciling utilitarianism and egoism. Unlike many Sidgwick scholars, I am not concerned with the DoPR in this thesis. I follow Sidgwick in believing his reconciliation between deontological and utilitarian principles to be separate from the problem of egoism, and therefore do not consider egoism to be an obstacle to the arguments I will be making here.

⁶³ Some important evolutions can be seen in the preface to the second edition. Sidgwick states that he has substituted ‘well-being’ with ‘happiness’, and that in his exposition of the utilitarian principle, he has decided to omit the last four words from Bentham’s phrase ‘the greatest good for the greatest number’ (viii) – both points are indicative of his understanding of utilitarianism. He has also revised III:XIII, on Philosophical Intuitionism, to give *more* of his own views, rather than commenting on those of other moralists, which is significant for my argument that Sidgwick did actually offer his own normative moral theory, despite his originally intended neutrality.

Related to the importance of the prefaces is the question of why, if the PD was this important to understanding his work, Sidgwick did not publish that document earlier. There is, as early as 1862, evidence from another source that Sidgwick was actually *intending* on a reconciliation in *ME*: In a letter to his great friend Graham Dakyns, he stated that he was ‘revolving a Theory of Ethics’, and ‘working on a ‘reconciliation between the moral sense and utilitarian theories’ (see also Arthur Sidgwick and Eleanor Mildred, *Henry Sidgwick: A Memoir* (London: Macmillan:1906). This, along with the PD, might also suggest the question as to why Sidgwick began *ME* from a neutral position at all, when he already knew that he would be drawing this particular conclusion in the book. To this I argue that *ME* does begin neutrally, and that Sidgwick is right when he re-emphasises in the preface to the second edition (presumably in response to Bradley) that he was not aiming to defend any one position over another. It is simply that in expounding the deontological and utilitarian positions in the way he describes in the PD, he had realised that there was a fundamental connection, and making this point

particular has prevailed into the present day. This is of course the enduring idea that Sidgwick is simply a classical utilitarian. It is on this point that account can be taken of Sidgwick's place in the history of moral philosophy itself.

0.7.b Sidgwick as a Utilitarian in the 20th Century

Despite Sidgwick's insistence that *ME* was not a defence of utilitarianism, it was received as such – and would continue to be applied as such⁶⁴. From the 1960s to the 1980s, for example, Sidgwick's position was utilised to provide support for the 'rule utilitarianism' that was developed in this time, most notably by Richard Brandt and R.M. Hare, in response to what were considered to be utilitarianism's less palatable aspects⁶⁵. In his development of his influential preference utilitarianism, or universal prescriptivism, Hare claimed to have combined 'the merits of both varieties' and he cites Sidgwick as being a particular type of utilitarian whose work provides support for those arguments⁶⁶. Given that Hare argues for a moral system in which the Kantian maxim of universalizability produces a form of (preference) utilitarianism, his work is discussed more fully below, in terms of its relationship to my similar

was a vital part of his investigation into those methods – but the whole of that investigation was still necessary in order for Sidgwick to make his point. The letter to Dakyns, however, does suggest something more of a specific intent regarding developing that reconciliation in *ME* – and this can only support my theory that Sidgwick saw and argued for that reconciliation.

⁶⁴ See chapter 1 (the literature review) for a more comprehensive overview of how Sidgwick's work has been read throughout the years.

⁶⁵ Richard Brant, *Ethical Theory* (Englewood Cliffs, N.J: Prentice-Hall: 1959), and 'A Theory of the Good and the Right'; in, *Philosophical Studies*, Vol.42 (1982) 119-139; R.M. Hare, *Moral Thinking* (Oxford: Clarendon Press: 1981).

Utilitarianism did indeed enjoy a period of greater credibility after the appearance of Brandt's 'rule' version. As will be seen below Hare's formulation of a version of rule utilitarianism was also particularly convincing. Rule utilitarianism also continues to be soundly defended today, most notably by Brad Hooker whose 2002 volume *Ideal Code, Real World: A Rule Consequentialist Theory of Morality* (Oxford: Clarendon Press: 2002) provides the paradigmatic defence of the theory. (See also Hooker, 'Rule-Consequentialism, Incoherence, Fairness', in, *Proceedings of the Aristotelian Society*, New Series, Vol.95 (1995) 19-35; 'Ross-style Pluralism versus Rule-Consequentialism', in, *Mind*, Vol.105, No.420 (1996), 531-552, and, 'Rule Consequentialism', in, *Mind*, New Series, Vol.99, No.393 (1990) 67-77). It was, however, initially attacked by critics (see J.J.C. Smart, 'Extreme and Restricted Utilitarianism', in, *The Philosophical Quarterly*, Vol.6, No.25 (1950) 344-354; Smart, *An Outline of a System of Utilitarian Ethics* (Carlton: Melbourne University Press: 1961); J.J.C. Smart and Bernard Williams, *Utilitarianism; For and Against* (Cambridge: Cambridge University Press:1973); and David Lyons, *Forms and Limits of Utilitarianism* (Oxford: Oxford University Press:1965) on the basis that 'rule' utilitarianism is simply not a viable form of the doctrine. This debate – and Sidgwick's place in it - is covered in chapter 1.

⁶⁶ Hare, *Moral Thinking*, p.43

argument built from Sidgwick's work in this thesis. Here, it suffices to note that Sidgwick was for this time highly relevant to the work of prominent utilitarians, and, ironically, that it was apparently Sidgwick's recognition of the *inadequacies* of utilitarianism (as opposed to Bentham and Mill, who did not provide a full critique of the doctrine) that provided a more acceptable place from which to start building a revitalisation of the utilitarian doctrine. The potential of Sidgwick's own solution – in which utilitarianism might be made 'credible' by making it dependent on deontological properties in a way similar to Hare's – was perhaps missed for this reason.

At a similar time to his appearance in Hare's work, Sidgwick was also afforded a meaningful place in the work of John Rawls, in Rawls' seminal and highly neo-Kantian work, *A Theory of Justice*⁶⁷. As Rawls was concerned to construct a form of the social contract that emphasised Kant's emphasis on intrinsic rightness, Rawls' use of Sidgwick is particularly illuminating of the peculiarly flexible way in which Sidgwick's arguments were perceived. Rawls recognised the importance of the utilitarian doctrine *only* as it had been expounded by Sidgwick, and Sidgwick thus retained his strongly utilitarian reputation – but Rawls *accepted* Sidgwick's version as the only really serious alternative to his own theory of justice as fairness⁶⁸.

There was then, at this point in history, an emergent sense that Sidgwick's work could be variously applied⁶⁹. Yet Rawls had still posited Sidgwick as a utilitarian. Sidgwick's utilitarian reputation survived unscathed – and was actually now on the rise. Hare was directly succeeded in his doctrine of preference utilitarianism by his student Peter Singer, and Singer, who today uses preference utilitarianism in a wide variety of ethical matters from bioethics to veganism (and often with great controversy) references Sidgwick liberally. Of particular importance to Singer's utilisation of Sidgwick is his 'point of view of the universe' passage (see III:XIII:382), which Singer takes to be representative of Sidgwick's entire moral view⁷⁰. For

⁶⁷ Harvard: Harvard University Press: 1971

⁶⁸ See *Theory of Justice*, p.417, and John Rawls, 'Foreword to 'The Methods of Ethics', in, Henry Sidgwick, *The Methods of Ethics* (Indianapolis: Hackett: 1981) pp. v-vi.

⁶⁹ Bart Schultz himself declares his own book, *Henry Sidgwick: Eye of the Universe*, to have come about as a direct result of this apparent flexibility in Sidgwick's work. See below.

⁷⁰ See most of Singer's major works – *Practical Ethics* (Cambridge: Cambridge University Press: 1979), *How Are We To Live?* (Oxford: Oxford University Press: 1997), Peter Singer and Katarzyna de Lazari-Radek, *The Point of View of the Universe: Sidgwick and Contemporary Ethics* (Oxford: Oxford University Press: 2014), and Helga Kuhse (ed.) *Unsanctifying Human Life* (Oxford: Blackwell Publishers Ltd: 2002), along with his numerous papers, articles, and essays.

Singer, Sidgwick appears in the capacity of an act utilitarian and, perhaps most controversially for the modern day - as a hedonistic utilitarian⁷¹.

These are not the only theorists who have referenced or utilised Sidgwick, although these are the most significant in terms of revelatory advances made in the field of utilitarianism. A more extensive account of the field of Sidgwick studies is given in the literature review. The point I wish to make with this outline here is that Sidgwick *has* maintained a quiet but steady presence in moral philosophy in the last hundred years, but that the real significance of Sidgwick for moral philosophy – for his potential to transcend the antithesis between deontology and utilitarianism that was going to go on to characterise so much of the debate in moral philosophy - has largely been missed, due to his being interpreted as a utilitarian.

0.7.c Recognition of Sidgwick's Reconciliation

Following the point that Sidgwick is almost universally considered to be a utilitarian, it must also be acknowledged that there has also been some movement towards recognising that Sidgwick perceived there to be not as stark a difference between utilitarian and deontological moral properties as his historical utilitarian status would suggest. This is also of course highly important to where this thesis fits in the field.

J.B. Schneewind, and Bart Schultz, in their respective books *Sidgwick's Ethics and Victorian Moral Philosophy* (1977) and *Henry Sidgwick: Eye of the Universe* (2004), have produced two works of major historical significance for Sidgwick. Any new interest in Sidgwick cannot fail but be enormously indebted to the incredible attention that is paid to the details of Sidgwick's life and work in both of these books,

⁷¹ Hedonistic utilitarianism, most commonly associated with Bentham (although Mill was also technically of this school) is thought to represent all of the very crudest and worst elements of the utilitarian doctrine. Hedonism itself is widely disparaged in general today, whether interpreted as the school of thought that holds pleasure to be the the most important intrinsic good, or whether used in non-academic terms to describe the behaviour of people who pursue pleasure at the expense of other, 'worthier' ends. I will not actually be disputing here the hedonistic element in Sidgwick's utilitarianism –although I do argue that he establishes the hedonistic principle far more convincingly than Bentham. I do, however, wish to dispute Singer's having attached such a direct label of hedonistic utilitarian to Sidgwick. Sidgwick should not, I believe, be primarily known as a utilitarian, and especially not as *this* kind of utilitarian (see my argument in chapter 3 for Sidgwick's rejection of pleasure as an ultimate end). Singer's work has been invaluable for interest in Sidgwick, and his regard for Sidgwick's work is always evident – but I do maintain that the reputation of hedonistic utilitarian is detrimental to a full understanding of Sidgwick as a moral philosopher.

and between them they have widened the field of Sidgwick studies phenomenally. Most significantly for my purposes here, because both Schneewind and Schultz are also acutely aware of Sidgwick's concern of the conflicting demands reason makes on morality. Schneewind pointed out in light of this that Sidgwick should not be accepted as a mere utilitarian. 'It is a mistake to view the book as primarily a defence of utilitarianism', Schneewind says, '...it is true, of course, that a way of supporting utilitarianism is worked out in detail in *The Methods*, and that there are places in it where Sidgwick seems to be saying quite plainly that utilitarianism is the best available ethical theory. From his other writings we know also that he thinks of himself as committed to utilitarianism, and that he assumes it in analysing specific moral and political issues. Yet it does not follow that the *Methods* itself should be taken simply as an argument for that position. We must try to understand it in a way that makes sense of its author's own explicit account of it'⁷².

Schultz specifically claims *Eye of the Universe* to have come about from an investigation into how it was that Sidgwick could feature in the conflicting arguments of the neo-Kantians such as John Rawls on one hand and the utilitarians such as Derek Parfit and Peter Singer on the other⁷³. Schultz is also clear, along with Schneewind, that Sidgwick is all too often read against a simple backdrop of classical utilitarianism, when this is not the most accurate understanding of his work. If Sidgwick was a utilitarian, Schultz says, "the system of utility takes in his hands a form so much more refined and delicate than was given to it by Bentham and James Mill, and is expounded with so many qualifications unknown to them, that it has become a very different thing, and is scarcely, if at all, assailable by the arguments which moralists of the idealistic type have brought against the older tradition"⁷⁴.

Awareness of this aspect of Sidgwick's work has been growing. Mariko Nakano-Okuno has recognised Sidgwick's systemising of Kantian principles with utilitarian ones⁷⁵; Janice Daurio reads ME as one moral theory⁷⁶; Brad Hooker recognises that Sidgwick was 'clearly looking for a moral system that would actually resolve uncertainty and disagreement', and theorises that Sidgwick's moral theory can be

⁷² Schneewind, *Sidgwick's Ethics*, p.192

⁷³ Schultz, *Eye of the Universe*, pp.1-20

⁷⁴ *ibid.* p.14

⁷⁵ See Nakano-Okuno, 'Sidgwick and Kant'.

⁷⁶ Janice Daurio, 'Sidgwick on Moral Theories and Common Sense Morality, in, *History of Philosophy Quarterly* (1997), Vol.14, No.4, 425-445

understood as a form of coherentism⁷⁷; Robert Audi includes Sidgwick as part of his recent defence of Rossian intuitionism, and uses the language of synthesis in reference to Sidgwick's concern with how we might be able to identify and rectify errors in moral judgments⁷⁸; Antony Skelton recognises that Sidgwick believed utilitarianism to be in need of a fundamental intuition⁷⁹; despite presenting Sidgwick as a hedonistic utilitarian, Singer and de Lazari-Radek do actually, in their own moral theory, defend *both* moral objectivism and hedonistic utilitarianism – and Sidgwick is relevant to both positions⁸⁰. Perhaps most notably, Sidgwick has appeared in the work of eminent scholar Derek Parfit, who himself specifically advances a reconciliation between deontological and utilitarian schools of thought⁸¹. All of these positions are important, and certainly contribute to making Sidgwick relevant not just to utilitarianism, but to the tension between utilitarianism and deontology. During the engagement with those theorists in the literature review, I argue against some theorists' utilitarian interpretations of Sidgwick, and in other cases envisage that the reconciliation in Rational Benevolence could compliment current similar theories about Sidgwick's reconciliation, or expand upon them. But I will ultimately argue that none of these positions fully draw out Sidgwick's full reconciliation, I as will try to do here. None have recognised Rational Benevolence as a full moral theory, nor argued for the immediate synthesis between deontological and utilitarian properties that it offers. None have argued that Sidgwick's utilitarianism – if it is to be called that – depends so extensively on a deontological principle that there could be grounds to argue for Sidgwick as a deontologist in some sense. These readings of Sidgwick, those that I aim to develop in this thesis, will be, I believe, the 'author's own explicit account' of utilitarianism to which Schneewind was referring⁸². Overall, I hope that

⁷⁷ Hooker, 'Sidgwick and Common Sense Morality', in, *Utilitas*, Vol.12, No.3 (2000) 347-360

⁷⁸ Audi, *The Good in the Right*, p.8

⁷⁹ Skelton, 'Henry Sidgwick's Moral Epistemology', in, *Journal of the History of Philosophy*, Vol.48, No.4 (2010) 491-519

⁸⁰ Singer and Lazari Radek, *Point of View of the Universe*, p.214

⁸¹ As was the case for Hare, where Sidgwick himself is relevant to Parfit's own moral theory is discussed in the literature review. Parfit's theory itself, however, clearly offers something similar to Rational Benevolence, and as such it is discussed below, in terms of its capacity as a potential rival or challenge to the place of Rational Benevolence in the field.

⁸² Sidgwick himself did not expressly dispute his utilitarian label, and it would therefore be presumptuous to do so here. I do maintain however that Rational Benevolence is *not* simply utilitarianism, and that for this reason it would be better to refer to it as Rational Benevolence, and not as Sidgwick's utilitarianism.

this thesis will contribute significantly to the existing but nascent inclination towards believing Sidgwick to have assimilated deontological and utilitarian principles.

0.8. Rational Benevolence in the Field of Modern Moral Philosophy

I have now outlined my argument that Sidgwick saw an essential and mutually informing relationship between deontological and utilitarian properties, and I have given the grounds for this argument as they appear (to me) in *ME*, and the PD. I argued that overall, Sidgwick's moral theory can be understood by the term 'Rational Benevolence': This term effectively summarises the rationality of both deontological and utilitarian principles, and the fact that the Universal Benevolence, or the utilitarian aspect, of the theory's title is only established on the basis of Sidgwick's understanding of the connection between ultimate rationality and rightness, which generates an unconditional and categorical 'ought'. I have also presented a brief overview of how Sidgwick has been variously read throughout the years, and thus largely framed the place for the arguments I will be presenting here. But there is a further aspect required in order to complete the understanding of this present work's place, which is a consideration of the potential applications, and limitations, of Rational Benevolence itself. I will first posit how I believe Rational Benevolence as a moral theory to be able to answer the problems that were outlined at the start of this introduction. I will then assess how Rational Benevolence fits into the field of moral philosophy in general.

0.8.a. Rational Benevolence and Moral Ambivalence: Descriptive and Normative Value

At the beginning of the introduction, I described –via the Parental Predicament - difficult ethical situations in which two very different courses of action appeared to be morally reasonable, and therefore morally justifiable, at the same time. Recognition of this dual reasonableness was seen to lead to the recognition that both courses of action were also morally reprehensible, precisely on the basis of the moral values of the alternative view. The result is serious moral ambivalence, which I then argued further is actually caused by the fact that the two types of moral value being represented here – absolute-based and outcome-based – are both reasonable, by

which it is generally meant that they are both somehow unconditionally right, and yet appear to be mutually exclusive on the basis of that reasonableness. It was then seen that a traditional 'divide' between deontological and utilitarian understandings of morality has permeated the level of moral philosophy, and that this translates into a failure on the part of individual autonomy to offer anything of real value to defending parents who are in the Parental Predicament.

I then suggested that the best approach to the situation would not be to better defend absolute-principle values and outcome based values separately (as this would offer nothing different to current approaches), but to explain the two moral properties as mutually informing. This reconciliation, I theorise, should take the form of a synthesis, wherein both properties retain their own independent value, and yet also rely on validation from the other, in a circularly interdependent relationship. Not only is this, I argue, a more holistic and realistic interpretation of the human moral experience, but explaining how the values of each moral property informs the values of the other is the stronger way to defend both reasonable courses of action in situations of serious moral ambivalence.

I argue that Sidgwick's theory of Rational Benevolence provides this very understanding of a synthesis between deontological and utilitarian properties that I have just described. Rational Benevolence shows that both types of moral property are ultimately *rational*, which directly accounts for my observation (and Sidgwick's) that both types of moral consideration as they appear in lived experience are reasonable. More comprehensively, in Rational Benevolence a deontological, non-consequentialist form of moral obligation (or duty) yields the utilitarian principle of Benevolence, which then itself also becomes a non-consequentialist moral duty, while also validating and confirming the roles of the principles that have constituted that duty. Rational Benevolence thus *explains*, or *describes*, the process of the human moral experience that is demonstrated by moral ambivalence, in such a way that neither one type of value or the other is discounted, but neither principle by itself is adequate for a full understanding of morality. Where I examine the potential of Rational Benevolence as a moral theory in full in the conclusion, I will refer to this theoretical-explanative function as Rational Benevolence's *descriptive* value.

I believe that the descriptive value of Rational Benevolence is important in itself for creating a more comprehensive and sympathetic understanding of the cause of moral ambivalence. But it will be recalled I said above that whereas mutual

limitation between two reasonable positions occurs more obviously, mutual validation would have to be proven – and it is my view that this is precisely what is being indicated by the descriptive function of Rational Benevolence. Rational Benevolence gives us the means of understanding at a philosophical level how the values of one position (non-consequentialism) actually necessarily take into account the values of the other position (utilitarianism), and vice versa, because in Rational Benevolence the two properties *depend* on each other - and it is this that I speculate could represent the *normative* potential of Rational Benevolence to be applied in practice. That is, actions could be justified morally if they could be seen to be a result of considering both the absolute-principle values and the outcome based values relevant to that particular situation in this interdependent way, i.e. if they could be seen to be *rationaly benevolent*⁸³. By the same line of reasoning, actions could also be found to be non-rationaly benevolent, if either one or the other value is completely absent. That is, as much as Rational Benevolence could provide a better means of justifying courses of action, it could also equally provide an effective check on certain courses of action⁸⁴. In this way, Sidgwick's theory of Rational Benevolence could perhaps offer to the framework of individual autonomy, or to the relevant healthcare policy, a form of moral philosophy that more robustly supports the difficult experience of autonomy (by better understanding its cause), and also a practical means of being able to defend certain courses of action, and identify appropriate limits to others.

0.8.b. Limits to Rational Benevolence as a Moral Theory

Although it is my view that Rational Benevolence offers a defensible moral theory in the way just outlined in the previous section, there are some areas in which its

⁸³ This appears to emerge as a very narrow form of reflective equilibrium. Although I do not argue this point specifically in the thesis, Sidgwick's relevance to reflective equilibrium features in the literature review, and there is certainly potential for Rational Benevolence to itself be explored from this point of view. There is also scope for reflective equilibrium to itself be considered as a form of reconciliation between deontological and utilitarian principles, but I argue that this is more of a practical amalgamation, and not necessarily a meta-ethical one.

⁸⁴ In the thesis conclusion this theory will be tested via examples such as parents who are determined to terminate a pregnancy on the basis of eye colour, and parents who are determined to continue a pregnancy when it is certain that the child will be affected by a painful, debilitating, life-limiting and usually fatal condition, such as Tay Sachs disease.

success could be called into question. This notion of 'success' can be broken down into two levels – the practical and the theoretical. It is my aim here to identify these challenges, and to formulate a response to them in such a way that they are acknowledged, but with Rational Benevolence still retaining a place in the field. This also includes a discussion of what I consider to be Rational Benevolence's main competitive theories. I will begin with the practical challenges and problems, before addressing the objections that might be brought against the actual theory itself.

0.8.b.i. Practical Limitations

The first objection that might be made is that Rational Benevolence, if it is successful in defending two courses of action as morally reasonable, simply justifies a situation that already exists, regardless of any moral philosophy. That is, it could be argued that most decisions made in the context of the Parental Predicament – and many other difficult ethical dilemmas - will *already* be the result of careful and painful reflection upon two apparent types of value, and that Rational Benevolence is not offering anything new by describing this fact. I admit that ethical decisions *are* likely to be of that nature at the private level, but the public opinion factor of the Parental Predicament demonstrates that parents are still often subjected to social influences, that arise as a direct result of there being two very difficult but defensible courses of action. I theorise that explaining the cause more robustly, in the way that I argue an actual synthesis between deontological and utilitarian moral properties to do, may go some way towards *protecting* those decisions, however hard they remain at the personal level.

Immediately following on from this is the claim that Rational Benevolence could never *so/ve* ambivalence. I entirely agree with this position. It is not my intention to 'solve' ambivalence; moral ambivalence is an essential part of human life, and is in many ways crucial to ethics in that it demands of us careful, involved, and considered responses to very difficult situations. Ambivalence does not apply to all people at all times – it is of course possible to hold definite moral views on all possible issues, prenatal decision making included. But where two conflicting views meet, with one propounding one set of values and the other another set of values,

this situation is reflecting the diverse range of possible approaches, and this is important. But this is also, I believe, where the problems start when it comes to real situations in which action of some kind must physically be taken – where a decision must actually be made. It is therefore my intention to explain the *cause* of ambivalence in the way that Rational Benevolence does, as I believe the recognition of the dual reasonableness of the two positions and the connection between them can be the only the key to providing autonomy with a more robust way of defending each course of action.

On the subject of Rational Benevolence's applicability to autonomy however, there is the related problem of implementation i.e. *how*, exactly, Rational Benevolence is meant to actually be applied in practice. I have stated above that it could be brought in as an informant of the relevant healthcare policy, but in reality there will eventually be contact between individuals and those who work under these policies (such as doctors and genetic counsellors). Rational Benevolence is a theory of moral philosophy, and it would most likely have to be modified in various ways, in order to assimilate it with policy in such a way that it is actually workable in some real sense.

Then there is the question of what rationally benevolent actions are to actually *look like*. This leads now to the first of the major theoretical objections that could be brought against Rational Benevolence. This is the question of how the Reason inherent in Rational Benevolence could possibly generate two different 'oughts'.

0.8.b.ii Theoretical Limitations: Rational Benevolence, Conflicts between 'oughts', and R.M. Hare

It was theorised above that moral ambivalence is due to two different positions on a matter of both being reasonable, and I have also outlined how Rational Benevolence specifically accounts for this dual reasonableness, which is what makes the theory relevant to this thesis. However, Rational Benevolence achieves this on the basis that rationality imposes an ought, and that as both deontological and utilitarian properties are rational, they must therefore both incur some kind of ought (which secures their moral value). There is an obvious problem here. If *rationality* is what imposes an ought, then how can it be said that *two* polar courses of action are both

‘oughts’?⁸⁵ For this is completely irrational – and actually brings us straight back to the problem of moral ambivalence from which we started: One cannot believe that one ought to do X *and* that one ought to do Y, because clearly the acknowledgement that it is right to do something involves the acknowledgement that it would be wrong *not* to do it. The two ‘oughts’ thus preclude each other, and therefore neither ought is tenable – and Rational Benevolence has not yet been seen to solve this problem in any practical way, simply by introducing the idea of ‘rationally benevolent’ actions.

R.M. Hare begins the development of his own moral theory in *Moral Thinking* from this very issue of the conflict between apparent oughts. Hare is a highly important point of comparison for Rational Benevolence. He specifically addresses moral conflicts from a point of view very similar to that which I made at the beginning of this introduction (and not dissimilar to Sidgwick’s), and proposes how they might be resolved (in a way dissimilar to mine). In doing so, Hare also provides one of the most effective and convincing reconciliations between deontological (specifically Kantian) and utilitarian principles that I believe to have been made to date, in the form of his preference utilitarianism. Here I give an overview of Hare’s moral theory, in order to examine his approach to moral conflicts, and subsequently the way in which he brings Kantianism into alignment with utilitarianism. I will then compare Hare’s solution to the system offered by Rational Benevolence, and address the problem of the apparent ‘two oughts’ in light of this.

Moral conflicts occur, according to Hare, in situations where we seem to have conflicting duties⁸⁶. There are many situations in which we clearly ‘ought’ to do two things, and we are clearly right to think that we ought to do both. But to decide that one ‘ought’ to do X involves the decision ‘*in some sense*’ that one ought not to do Y⁸⁷. This leads to a situation in which one ought to do x and also ought not to do x, which contradicts the claim that ‘ought’ means ‘can’, because clearly it is impossible to both do something and not do it⁸⁸.

⁸⁵ It is precisely this that Sidgwick’s system allows for, as I shall try to argue throughout the thesis overall.

⁸⁶ Hare, *Moral Thinking*, p.26. See also Hare, *Sorting out Ethics* (Oxford: Clarendon Press: 2000).

⁸⁷ *Ibid*, p.27

⁸⁸ Hare’s observation of moral ambivalence is clearly close to mine – but Hare does not point to the apparent reasonableness that I have argued here causes moral conflict. This is because Hare proposes a very different solution to these conflicts, as will be seen. I have also avoided using ‘ought’ in the context of the Parental Predicament, and I think this is the right thing to do in situations of grave moral ambivalence, such as beginning of life decisions, or end of life decisions. It is very difficult to say someone ‘ought’ to end a pregnancy, or the life of someone for whom withdrawal of treatment or

Hare believes this moral conflict to arise from the fact that there are two levels of thinking in the human moral experience – the intuitive, and the critical. As most moral conflicts are the result of conflicting intuitions (or conflicting ‘oughts’), Hare’s solution to moral conflict – and the beginning of his normative theory – is his claim that real moral rationality takes place at the critical level. This philosophical logic grounds moral reasoning entirely in the rational, and removes it from intuitionism⁸⁹.

At the critical level of thinking, it becomes possible to see the logical properties of moral language - that is, what we must really *mean* when we speak of moral obligation. These logical properties of moral words are prescriptivity of moral judgements, overridingness, and universalizability. This condition that moral principles be truly universalisable is Hare’s true innovation in the field of utilitarianism. Universalizability requires that we are prepared to prescribe the same moral judgement for any other precisely similar situation, from our position of critical thinking. As the thinker could be in the position of any other party in the situation, he must make an impartial judgement that ‘does best’ for all the parties involved. This thinking must be rational, and based on the facts of a situation; the question of what is best is itself rational, and therefore the answers must be actual truths⁹⁰. The only facts we can know of are those pertaining to the position of each individual in the situation to which the moral principle is being prescribed – and these facts are the individual’s preferences⁹¹. Through this process, impartial universal prescriptivism leads directly, rationally, and logically, to preference utilitarianism. An alternative and perhaps more accurate way to state this is that preference utilitarianism is logically

euthanasia has for some reason become relevant. On the other hand, Sidgwick’s theory of Rational Benevolence depends upon a crucial relationship between ‘right’ and ‘rational’, that is necessary for giving the theory a rational ‘ought’. This contradiction and its proposed solution are discussed below.

⁸⁹ The intuitive level, Hare argues, is insufficient. Hare does not reject our basic moral inclinations entirely; he recognises that we are taught general principles from childhood, that it is important that this is the case, and that they do have some role (*Moral Thinking*, pp. 30-39) for moral learning, teaching, and for when there is little time for critical thinking. Hare actually borrows Ross’s term ‘prima facie principles’ to represent the principles that our intuitive level of thinking presents. But this one-level moral thinking is not completely adequate. These principles have no claim to authority beyond the fact that the individuals who hold them *do* hold them, and this may well have been affected by influences of any kind – prejudice, for example (*Moral Thinking*, p.12, p. 40). There is nothing to which these intuitions can refer in order to make them truly self-justifying. Hare’s thought here bears more than a passing resemblance to Sidgwick’s argument that common sense morality principles are important, and valid to some extent, but are in need of rational synthesis.

⁹⁰ *Moral Thinking*, pp.87-89

⁹¹ *Ibid*, pp.92-106.

derived from universalisable prescriptivism. This is how Hare's system completely incorporates Kant's universalising maxim – and produces the same results⁹².

Hare draws attention to this very point about the similarities between the results of his preference utilitarianism, and the Kantian system. He is among the highest profile philosophers to have recognised that Kant and the utilitarians have been 'at loggerheads' – but also that utilitarian and deontological principles *both* have 'hold of a part of the truth'⁹³. He does not believe there to be a great separation between Bentham's statement "everybody to count for one, and nobody for more than one" and Kant's 'act only on that maxim through which you can at the same time will that it should become a universal law'⁹⁴ – and the combination of universal prescriptivism with the preferences of all parties in his own moral theory clearly picks up, expands upon, and unites both of those maxims, with startling clarity. Whereas Hare states that we owe the practice of studying the formal properties of moral argument above all to Kant, the conclusions of Hare's particular moral reasoning 'yield a system...whose conclusions have a content identical with that of a certain kind of utilitarianism'⁹⁵.

There are several ways in which Rational Benevolence needs to be compared with Hare's system, in order to show where still Rational Benevolence has relevance in the field alongside this masterful and innovative type of reconciliation⁹⁶. Firstly,

⁹² Despite his firmly utilitarian position, Hare never denied that his study into the logic of imperatives had been inspired by Kant. He was also clear that Kant's line of reasoning, 'yields a system...whose conclusions have a content identical with ...utilitarianism' (*Moral Thinking*, p.4). It is widely argued that *Sorting out Ethics* represents Hare as being as much of a Kantian as he was a Utilitarian – indeed, the book contains a section entitled 'Could Kant Have Been a Utilitarian?' But Hare does depart significantly from Kant, even on the point of universalizability and the categorical imperative. It simply makes no sense, to Hare, to make 'do not steal' a categorical maxim, and doing so would result in absurdities, such as preventing us from being ethically justified in stealing nuclear weapons away from a despot (*Moral Thinking*, p.164).

⁹³ *Moral Thinking*, p.vi, 4

⁹⁴ *Ibid*, p.5

⁹⁵ *Ibid*, p.4

⁹⁶ Many scholars have recognised some kind of affinity between Hare and Sidgwick. Rather fittingly, Richard Brandt described Hare's preference utilitarianism as the most substantial formulation of utilitarianism since Sidgwick (*A Theory of the Good and the Right*, p.713). See also Tom Carson, 'Hare on Utilitarianism and Intuitive Morality', in, *Erkenntnis*, Vol.39, No.3 (1993) 305-331; Tim Mulgan, *The Demands of Consequentialism*, (Oxford: Oxford University Press: 2005); Geoffery Scarre, *Utilitarianism* (London: Routledge: 1996); Roger Crisp, *The Cosmos of Duty* (Oxford: Oxford University Press: 2015); Robert Shaver, 'Sidgwick's Minimal Metaethics', in, *Utilitas*, Vol.12, No.3 (2000) 261, and 'Sidgwick on Moral Motivation', in, *Philosopher's Imprint*, Vol.6, No.1 (2006) 1-14; and Peter Singer, *Practical Ethics*. Schultz specifically states that Hare 'appropriates nearly all of Kant's ethics for his own purposes', and that in doing so he is a direct echo of Sidgwick (Schultz, *Eye of the Universe*, p.195). Mariko Nakano-Okuno believes Sidgwick to have clarified in more detail than Hare the relationship between utilitarianism and Kantian ethics (Nakano-Okuno, *Sidgwick and Kant*, p.262). Daniel Jacobson points out that Sidgwick anticipated the idea of an esoteric, multi-level

there is a striking similarity between Sidgwick's process of establishing of Universal Benevolence and Hare's eventual procurement of preference utilitarianism. Sidgwick's is a broader version (Hare's language of preferences is more specific), but he began from a Kantian universalisation maxim (Justice), from which he derived a utilitarian principle (Universal Benevolence). Secondly, Sidgwick's utilitarian principle is identified via the relation of integrant parts to each other which generates the Point of View of the Universe, and there is a clear parallel here with Hare's argument for all preferences⁹⁷. Also, both Hare and Sidgwick use a careful process of rationality to arrive at their moral precepts (Hare's emphasis on the logic of moral language is arguably more compact and direct) – and both, of course, present a form of reconciliation⁹⁸. There is little doubt that in many ways, Hare's theory – and his reconciliation - is very convincing. As Walter Sinnott-Armstrong points out, the logic of his argument is very hard to argue with⁹⁹. He has used logic and rationality to derive straight from simple moral language the vast ideas of universal prescriptivism, which then cogently lead to a place in which utilitarianism seems to be the *only* answer. And it fits perfectly with Kant's maxim. In fact, Hare makes this seem like a perfect extension of Kant's doctrine. For these reasons, I do not necessarily propound to be doing anything *better* than Hare. But I do argue that there are two crucial points of departure between my argument for Sidgwick's Rational Benevolence which allows this reconciliation to do something *different* to Hare's system. The first is that Sidgwick's reconciliation does not, on my view, produce straight-up utilitarianism, and this is more useful to the 'divide' than a reconciliation that still produces either

morality, that was then fully developed by Hare in his two-level utilitarianism (Jacobson, 'J.S. Mill and the Diversity of Utilitarianism', in, *Philosopher's Imprint*, Vol.3, No.2 (2003) p.2).

⁹⁷ Sidgwick's position, like Hare's, also incorporates both Bentham's and Kant's maxims.

⁹⁸ Hare is of course specific that he is aiming for a reconciliation between Kantian and utilitarian principles, where Sidgwick was not. Hare also specifically uses the language of 'synthesis' to refer to the amalgamation of Utilitarianism with Kant's doctrine (*Moral Thinking*, p.43; see also Robert D'Amico, 'R.M. Hare, 1919-2002', in, *Proceedings and Addresses of the American Philosophical Association*, Vol.76, No.2 (2002) 129-130) despite specifically denouncing deontology itself on the grounds that it contains no coherent, rational account of moral thought (Hare, *Rules of War and Moral Reasoning*, in, *Philosophy and Public Affairs*, Vol.1 No. 2 (1972) pp. 166 -181). Sidgwick would have agreed with Hare that common sense morality (which for Sidgwick represents the deontological position) by itself is inadequate as a full system of ethics. Furthermore, the only principle from that method that Sidgwick *could* find to be ultimately rational is Kant's maxim (which then establishes Benevolence, as Hare's universalising factor established preference utilitarianism). However, it is my view that Sidgwick is actually more sympathetic to the rules of common sense morality (see chapter 2 of this thesis), and that this is part of what gives a different route of reconciliation in Sidgwick, as opposed to Hare.

⁹⁹ Walter Sinnott-Armstrong, 'R.M. Hare, (1919-) <http://sites.duke.edu/wsa/papers/files/2011/05/wsa-rmhare2001.pdf> (accessed 18/6/2016) p.332.

exclusively deontology or utilitarianism. Hare's logic of prescriptivism 'uses' universalizability to generate utilitarianism as an end result, and Hare argues for utilitarianism as the only possible result of the logic of moral language, and prescriptivism. I admit that there are passages in *ME* in which utilitarianism appears to logically emerge as the superior theory, but I argue that all of these must be read in the context of Sidgwick's wider argument for the relationship between utilitarianism and deontology, and I will make these arguments in this thesis. I also argue that even if Sidgwick's theory *does* eventually have to be called utilitarian in a formal 'moral philosophic' sense, that the mutually and circularly informing relationship between deontological and utilitarian moral properties in Rational Benevolence would still be at an advantage for attempting to overcome the divide between them.

Furthermore, Rational Benevolence offers something practically and substantially different to Hare's, and this is crucial to the problem of moral ambivalence from which this thesis starts, and to the way in which I have suggested that ambivalence is handled. This is the potential to justify two courses of action in situations of serious moral ambivalence, rather than just one – which is precisely something that Hare rejects. On Hare's view, the critical level of thinking will reveal that only one course of action – only one 'ought' is truly right. In short, according to Hare, 'if your duties conflict, one of them is not your duty'¹⁰⁰. He has very good grounds for arguing this (taking into account preferences, which are factually and thus rationally based) and his system thus appears to overcome the conflict between duties by revealing only *one* or the other to be the actual duty. But this position cannot truly help the situation of moral ambivalence, which I have argued is itself indicating something of value – which is that *both* positions are reasonable. It is my overall purpose in this thesis to explain and justify ambivalence by proving via Sidgwick's synthesis how both are reasonable. But this does bring us back to the problem from which this section started, and which Hare seems to have so effectively addressed - which is that conflict between rational oughts is irrational. How are two reasonable i.e. rational courses of action to be held right, or rational, at once, when rationality and rightness on Sidgwick's view incurs a categorical moral obligation?

There are three possible ways of framing the outcome of this problem: 1) Both are 'oughts' in which case this is irrational; 2) Only one is an ought, thus cancelling

¹⁰⁰ *Moral Thinking*, p.26

out the other and defeating the object of addressing ambivalence; 3) Neither can be established by Rational Benevolence to be an ought, in which case the justification for them both as being reasonable collapses. In order to address moral ambivalence, and in order to produce a theory that has any hope of comparing to Hare's approach to moral conflicts, we need to begin from the acceptance of the equal reasonableness of two different types of principle. This means that only the first option – 'both are 'oughts'', is viable. This therefore requires proving, somehow, that this situation is actually not irrational.

The first part of this argument is to point out that I am actually *not* aiming to establish two independent 'oughts', or two independent 'duties'. The point of Rational Benevolence is exactly the opposite, in that I am specifically trying to find a way in which the dual 'oughts' present in the deontological and utilitarian properties of each position inform and limit each other, precisely so that two reasonable positions *can* exist at the same time. This is the entire point of this exercise – to account for *both* types of reasonableness, which I have argued is at the core of so much ethical conflict. This is what is offered by Rational Benevolence, and this is the major difference between Hare's treatment of moral ambivalence (such as it appears in *Moral Thinking*), and mine, or that of Rational Benevolence. It is Sidgwick's particular argument for *how* Reason instils rightness in what is ultimately rational that is the key to this dual reasonableness – and specifically the fact that Sidgwick believes rightness to lie both in ultimately rational ends, *and* in certain ultimately rational actions¹⁰¹. It is this shared appeal to Reason that Rational Benevolence represents: Dual, mutually informing reasonableness, in which the reasonableness of both is indivisible from the reasonableness of the other. That model, I argue, then translates to the practical level wherein 'rationally benevolent' actions embody that mutually informing rightness. This, then, is what 'rationally benevolent actions' actually look like. And they are specifically capable of defining more than one course of action.

It might be better to say, in light of this, that both courses of action are *right*, rather than 'oughts' – or even that both courses of action are 'morally justifiable'.

¹⁰¹ This means that we have to admit some kind of consequentialist version of moral obligation into the equation, alongside my emphasis on Sidgwick's deontology, but this of course is the point of my argument for Rational Benevolence. It will be seen in chapters 2 and 3 that Sidgwick includes the concept of ultimate rational ends as part of his argument for a non-consequentialist understanding of rightness, and then in chapter 4 that it is indeed the utilitarian principle that Sidgwick eventually establishes as the end of moral action. But whereas Benevolence may be consequentialist in its dictates, as an ultimately rational self-evident principle, it is, itself, non-consequentialist.

Rational *Benevolence itself* is the ought, the ultimately rational moral rule— then the mutually informing relationship between deontological and utilitarian properties can inform how the two related courses of action are *right*¹⁰². But this part of the discussion is highly normative, and it is not my primary aim in this thesis to establish a theory of normative ethics. This too, may be one of the theory's theoretical limitations. As a normative theory it is certainly at a disadvantage as compared to Hare's amalgamation of universalisation with preferences of interested parties - although I will, in the thesis conclusion, attempt to outline how universalizability might possibly apply to rationally benevolent actions.

I do maintain however that whereas Hare says that moral conflict is not resolvable, this is not necessarily the case. Instead, I believe that the problem needs to be reframed. The problem has lain in the attempt to prove one position over another (and this *will* be irresolvable, on my view, which is what has led to the perpetuation of the moral-philosophic divide), when actually the best approach may be to recognise that the solution is to adopt neither one nor the other, but *both*. There are two sources of moral rationality as Sidgwick proves – the way to handle this fact is to incorporate it, not erase it. Rational Benevolence establishes both moral properties as rational and right, in such a way that they no longer conflict, but rather are both

¹⁰² This formula raises another possible parallel between Rational Benevolence and Hare's work, which occurs regarding act/rule utilitarianism. The 'certain kind of utilitarianism' to which Hares refers (*Moral Thinking*, p.4) is one that is act-utilitarian at the critical level of thinking (because it leaves out no feature of an act which might be relevant), and rule-utilitarianism when the general prima facie principles used at the intuitive level are included (*Moral Thinking*, p.43). Hare believes the controversy between the two types to have arisen because the two different levels of thinking have been ignored – 'once the levels are distinguished, a form of utilitarianism becomes available which combines the merits of both varieties'. (Interestingly, it is the act utilitarian aspect that Hare states provides the means through which 'utilitarians and Kant get synthesised' (*Moral Thinking*, p. 43) – this is because of the combination of universalizability, and the consideration of all preferences of all relevant parties.) Given my argument that Rational Benevolence provides an overall framework through which 'rationally benevolent' actions could be identified, it might reasonably be asked whether Rational Benevolence is also a form of two-level utilitarianism, like Hare's. Before this question can be answered, it must be noted that Sidgwick's system is actually the reverse of Hare's. From intuitions, he has derived the general principle of Rational Benevolence, which could be called a general rule-utilitarianism. Rationally benevolent actions are those that would have to consider every feature of the situation on a case by case basis, and therefore represent what could then be thought of as act utilitarianism. I do not deny that Rational Benevolence appears to conform to this understanding of utilitarianism, but I do maintain that we can lift it out of rule/act utilitarianism altogether, because of the mutually informing nature of the relationship between them, in which utilitarianism depends on non-consequentialist deontological properties. I also wish to direct focus away from interpreting Sidgwick to simply be a utilitarian, whether that is of the act or rule variety. (Brad Hooker's 'Sidgwick and Common-sense Morality' p. 355-360 provides a particularly clear discussion of whether Sidgwick was an act or a rule utilitarian – see below. See also Hooker, Brad, 'Rule Consequentialism', The Stanford Encyclopedia of Philosophy (Winter 2015 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2015/entries/consequentialism-rule/>>. (date accessed: 24/6/2016).

essential and irrefutable components of the human moral experience. This, I hope, might only add to the field of work such as Hare's that argues for the affinity between deontology and utilitarianism.

0.8.b.iii Theoretical Limitations: Other Rival Theories

Although Hare's reconciliatory system is probably the closest in nature to Sidgwick's Rational Benevolence, there are three further writers whose moral theories also offer other theories that attempt reconciliation between moral principles¹⁰³. They are framed here in terms of the challenge that they might present to the place of Rational Benevolence, and how I envisage Rational Benevolence to still be relevant alongside these sophisticated moral theories.

0.8.b.iii a) W.D. Ross: Deontological Intuitionism

The next writer whose theory recognisably offers a combination of deontological and utilitarian principles is W.D. Ross. Ross was a deontological intuitionist, whose moral theory – mainly presented in *The Right and the Good* (1930) and *Foundations of Ethics* (1939) - specifically includes elements of the two concepts thought to represent the opposing priorities of utilitarian and non-utilitarian positions, 'good' and 'right' respectively¹⁰⁴. He is a true intuitionist, believing capturing the essence of our moral impulses to be more important than their systemisation, as, say, Henry Sidgwick would have argued¹⁰⁵. Ross rejected both ideal utilitarianism (the doctrine of G.E. Moore) because it holds that what makes a right act right is purely that it maximises good, and Kantianism universalizability, because it oversimplifies the reality of the moral life, and fails to take into account the salient fact that acts often have more than one motive (not simply the motive of what one thinks one 'ought' to do)¹⁰⁶. For Ross, 'common-sense morality' – the moral perceptions of ordinary,

¹⁰³ As I have mentioned above, the authors of these theories do actually reference Sidgwick – one of them extensively. These theories are therefore important to where this current work fits in the field of Sidgwick studies. However, as it is the remit of the literature review to assess how and where Sidgwick has been read, the particular relevance of *Sidgwick* to these theorists is covered there. Here, the theories are assessed in terms what they offer in comparison to Rational Benevolence

¹⁰⁴ See Ross, *The Right and the Good* (Oxford: Oxford University Press: 1930), and, *The Foundations of Ethics* (Oxford: Oxford University Press: 1939)

¹⁰⁵ See *The Right and the Good*, p.19; *Foundations*, p.83

¹⁰⁶ *The Right and the Good*, pp.19-20

thoughtful, educated individuals - is the main source of what he termed 'the data of ethics'. In basic terms, this 'data' is intuitively perceived principles that we consider to be right, or duties in some form. With wonderful directness, Ross points out that we do have these duties, but there are clearly cases in which the circumstances require that a certain duty is broken¹⁰⁷. In this way, Ross too – like Hare – deals with conflicts between moral duties. Using the example of breaking a promise to meet a friend in order to bring relief to victims of an accident, Ross' presents his innovative view of what is happening morally in this situation. Ross does not concede that we are merely looking to what will secure the best outcome – he maintains focus on the nature of the duties themselves, arguing that they *remain* duties, but in the case of promise breaking vs distress relieving, we are recognising that in those circumstance, relieving distress is at that time *more* of a duty¹⁰⁸. On Ross' view, this type of duty has a morally significant value, even though it is not a 'duty proper'; he refers to them as *prima facie* duties (or 'conditional duties')¹⁰⁹. That is, they can be overridden, and they can be overridden on the basis that in particular circumstances, where duties conflict, one duty will reveal itself to be more urgent than another¹¹⁰. In concluding what right acts are, Ross states, in parallel with my observation about dual reasonableness given at the beginning of this introduction, that every act is *prima facie* right and *prima facie* wrong, and the right acts are those which have the greatest balance of *prima facie* rightness¹¹¹. Whereas there is no fixed principle that can tell us how these intuitive duties are to operate/override each other, there are

¹⁰⁷ Ibid, p.19

¹⁰⁸ Ibid, p.18. Ross thus specifically rejects the idea that what is *right* is right because it is productive of the most good. But he also rejects Kant's view that duties admit of no exceptions in favour of 'imperfect obligations', such as relieving distress.

¹⁰⁹ Ross lists six divisions of duties that are revealed by our common moral convictions – duties of fidelity/reparation, duties of gratitude, duties of justice, duties of beneficence, and duties of self-improvement (*The Right and the Good*, p.71).

¹¹⁰ Ross states: "When I am in a situation, as perhaps I always am, in which more than one of these *prima facie* duties is incumbent on me, what I have to do is to study the situation as fully as I can until I form the considered opinion (it is never more) that in the circumstances one of them is more incumbent than any other; then I am bound to think that to do this *prima facie* duty is my duty *sans phrase* in the situation' (*The Right and the Good*, p.19). These duties are also duties at different times for different reasons, according to the circumstances (ibid, p.24). Ross argues that they are not actually different from Kant's absolute duties, that require us to do 'one particular act in particular circumstances' (ibid, p.28), and that we know this because in cases where one duty has overridden another (for example, if we have had to break a promise) we still recognise a *prima facie* duty to keep the promise, and we feel compunction for having not done so. Hare also recognises the role of compunction and remorse at the intuitive level (*Moral Thinking*, pp. 28-31), but does not afford it nearly so much moral significance as Ross.

¹¹¹ *The Right and the Good*, p.41, p.46

four things that are intrinsically good – justice, pleasure, knowledge and virtue - that inform our ‘all things considered’ decisions when deciding on prima facie duties¹¹².

Ross’ prima facie duties presented a pioneering idea. Instead of replicating the sorts of attempts to affix one primary principle to human morality, Ross’s distinct moral philosophy is to allow for both a variety of moral requirements, or obligations, *and* the presence of intrinsic goods. In Ross’s system both the right *and* the good are the result of independent intuitions: Moral requirements therefore cannot be reduced to some other ultimate, fundamental principle (such as the principle of utility, or testing by the categorical imperative), and *nor can* goods be found to be second to some ultimate good¹¹³. As Antony Skelton pointed out, such a system seems to escape the insufficiencies of utilitarianism, and the excesses of deontology¹¹⁴.

The Right and the Good is still held up as one of the most important works of twentieth century moral philosophy, and it accords well with many peoples’ experiences. But it has not gone without criticism. Ross himself predicted the likely objections. He admits that as a moral theory it is unsystematic, although it is not

¹¹² Ibid, p.134

¹¹³ Ibid, p.115, p.118; *Foundations*, p.19, p.180, p.278

¹¹⁴ Antony Skelton, ‘On Henry Sidgwick’s ‘My Station and its Duties’, in, *Ethics*, Vol.125 (2014) 586-591. It has since been maintained by many that Ross’s idea of *prima facie* duties is one of the most important contributions to modern moral philosophy (see C.D. Broad, ‘Critical Notice of W.D. Ross, *Foundations of Ethics*’, in, *Mind*, Vol.49 (1940) 228-239 ; A.C. Ewing, *Second Thoughts in Moral Philosophy* (New York: Macmillan:1959), and Phillip Stratton-Lake (ed.), *The Right and the Good* (Oxford: Oxford University Press: 2002) p. xxxvii), and that Ross specifically overwrote Sidgwick’s argument against self-evident moral principles in such a way that it did not simply collapse back into utilitarianism (see Lazari-Radek and Singer, *Point of View of the Universe*, p. 87), as Sidgwick’s was perceived to have done. It could then possibly be argued that Ross’s system does in fact achieve something very like that which this thesis claims only Sidgwick achieved, which is the uniting of a degree of absolute moral obligation with a natural consideration for outcomes. After all, Ross describes goodness as ‘a salient element in the ground of [the act’s] rightness’ (*Right and the Good*, p. 47). But the opinion on Ross’s success is by no means unanimous. Lazari-Radek and Singer, for example, argue that whereas Ross’s system overcame problems such as the firm Kantian limitation on lying when a life might be saved as a result of that lie, and did so without simply looking to pure consequences, it could not adequately meet Sidgwick’s challenge that some promises are made, or some truths told or kept erroneously, and as the result of inadequate thought. They draw on W.A. Pickard-Cambridge’s critique of Ross, in which Pickard-Cambridge presented a series of promises that had been made in such circumstances. Although Ross responded to each example, Lazari-Radek and Singer argue that he only showed how much room there was for disagreement, over precisely what these *prima facie* duties were to be (*Point of View of the Universe*: pp. 85-87). It is this confusion that is generally thought to considerably weaken Ross’s argument (see Shelley Kagan, *The Limits of Morality* (Oxford: Clarendon Press: 1989) p.17, specifically Kagan’s argument that Ross should perhaps have used the terminology of ‘*pro tanto*’, rather than ‘*prima facie*’). Whereas the general principles (from which Ross derives the *prima facie* duties) are meant to be self-evident, how we should go about deciding which duty is to be obeyed in which circumstance is not. Reflection, Ross says, is our best chance of acting rightly – but as Lazari-Radek and Singer also point out, Ross does not suggest how we might go about this reflection, and ultimately Ross is thrown back on a heavy reliance on simple intuition, rather than a robust normative or meta-ethic.

claiming to be anything else: It works primarily from self-evidence and intuitionism. But it is in that intuitionism that further weakness lies. Intuitionism was – and is – roundly criticised for almost always failing to provide any actual, solid answers, and for doing little more than telling us what we already know¹¹⁵. The problem in Ross' system has always been that there is little ground on which Ross can propound these duties as duties, and actually propound them as *right*¹¹⁶.

Like Ross, Sidgwick also aimed to give some kind of systemisation to a common sense morality, and his theory is of course also comprised of self-evidence and intuitionism. But given Sidgwick's heavy use of Reason to defend even his intuitionist principles, Ross emerges as more of a pure intuitionist than Sidgwick. This is where I believe Sidgwick's approach to the two moral properties can hold up where Ross' may not. Sidgwick's system specifically demonstrates that both deontological and utilitarian properties are *rational*, and establishes a logical connection between them on this basis. This rescues Sidgwick's epistemology from being entirely dependent on intuitionism, and grounds it in something more substantially objective. If we reject Ross' theory that we do just see these duties, then we must reject Ross altogether. Ross' system gives a plausible argument for a plurality of duties, but it is still just that – a plurality, of which we are already all too aware¹¹⁷.

0.8.b.iii.b) William Frankena: 'Mixed Deontological Theory'

¹¹⁵ See G.J. Warnock, *Contemporary Moral Philosophy* (New York: St. Martin's Press: 1967), p.13; W.D. Hudson, *Modern Moral Philosophy* (London: Macmillan:1970), p.105; Alasdair MacIntyre, *A Short History of Ethics* (London: Routledge: 1967) p.255; Richard Kraut, *What is Good and Why* (Cambridge, MA: Harvard University Press: 2007), p.123; David Wiggins, *A Sensible Subjectivism* (Oxford: Oxford University Press:1998).

¹¹⁶ H.W.B. Joseph made this point by drawing particular attention to Ross's lack of a single unifying principle (the sort for which moral philosophers from both schools of thought had previously been striving, and the sort that Prichard claimed did not exist), arguing that this leads simply to a collection of 'unrelated obligations', which is not an acceptable basis for a comprehensive moral theory (Joseph, *Some Moral Problems* (Oxford: Oxford University Press: 1931). Joseph argued that Ross's work in general simply was not systemised enough. In 1971, Rawls would make the same point, arguing like Joseph that without any 'reasonable ethical criteria' with which to weight normative ethical principles against each other, we have only half a theory (*A Theory of Justice*, p. 41)

¹¹⁷ Part of the appeal of Ross' theory is that he accounts so well for both *good* and *right*, those two concepts that had been so fundamentally split with Kant/Bentham, and later with Moore. Sidgwick does not use this language as specifically as Ross does, and actually does not pay much direct attention to the place of 'good' at all. There is, however, potential for Sidgwick's theory to be phrased in terms of the roles of 'right' and 'good', if these two concepts are equated with deontological and utilitarian moral properties respectively. Whether or not Rational Benevolence does combine these two concepts will be addressed in the thesis conclusion.

American philosopher William Frankena can, like Ross, also be thought of as holding a deontological intuitionist position. But Frankena was far more explicit than Ross that both deontological and utilitarian theories of morality contribute something vital to morality. Where Ross incorporated utilitarian ideas via a metaphysical argument for intrinsic goodness, Frankena specifically sought to give credit to the central utilitarian principle itself¹¹⁸. This is, according to Frankena, quite simply because “we must grant that the utilitarians have hold of an important part of the truth”¹¹⁹. The construction of his moral theory that demonstrates his deontological/utilitarian combination is as follows.

In a section from the 1973 reprint of Frankena’s *Ethics*, that Louis Pojman expressly entitles ‘a reconciliation of ethical theories’, Frankena presents the idea that for a theory of moral obligation, we first ‘recognise two basic principles of obligation, the principle of utility, and some principle of justice’¹²⁰. The principle of utility, he argues, must be first taken as ‘one of our basic premises’, because we can only make sense of morality by talking in terms of the maximising of good and the minimising of evil¹²¹. But for this reason the principle of utility itself is not, for Frankena, actually the most primary principle, simply because it only indicates the even more basic premise that we ought to maximise good and minimise evil. The actual primary premise that is indicated by the practical principle of utility is, then, the principle of beneficence. This principle in turn requires another that tells us how to distribute goods and evils, when there are conflicting claims being made. This is the principle of distributive justice, through which our categorical *duty* to maximise good and minimise evil is established and guided. This general theory, Frankena points out, ‘would be a deontological one, but it would be much closer to utilitarianism than

¹¹⁸ Frankena, *Ethics* (Englewood Cliffs, N.J.: Prentice-Hall, Inc.:1963) pp.29-46.

¹¹⁹ Hare’s directly echoed these words of Frankena’s, almost 20 years later, although he does not reference him.

¹²⁰ See Louis Pojman, *Ethical Theory: Classical and Contemporary Readings* (Belmont, CA: Wadsworth:1994) p.239

¹²¹ Frankena is exacting in his view of the role of utility, extending his argument to ultimately claim that all of our moral obligations and rules, even those of justice, presuppose the existence of good and evil, and the effect on our lives. “It is not easy to deny”, Frankena says, “as pure deontologists do, that one of the things we ought to do...is to bring about as much of a balance of good over evil as we can...I find it hard to believe that any action or rule can be right or wrong...in the moral sense, if there is no good or evil connected with it’ (*Ethics*, p.36). The ‘minimising of evil’ had developed its own field of utilitarianism at this time, under Karl Popper’s term ‘negative utilitarianism’. The focus in this branch of the doctrine is attempting to secure the least amount of harm, rather than the most amount of good (Karl Popper, *The Open Society and its Enemies* (Oxford: Routledge and Kegan Paul Ltd: 1966).

most deontological theories'; as Frankena suggests, we might call it a '*mixed deontological theory*'.

Frankena's system is arguably less innovative than Ross', and not as rigorously rational and systematic as Hare's. But his straightforward argument for the roles of deontologically-understood duty and the utilitarian principle has the effect of appearing to dissolve the conflict between them with ease¹²². In terms of the relationship between Frankena's mixed deontological theory and Rational Benevolence, Frankena's theory is similar to the early stages of Sidgwick's construction of Rational Benevolence, in which Sidgwick recognises a principle of justice that will form the basis of duty for the ultimate principle of Benevolence¹²³. As was the case with Ross' reliance on intuition, Frankena's argument for the connection between the duty of justice and the utilitarian principle of beneficence is lacking the sequence of rational steps through which Sidgwick arrives at his moral theory, and Sidgwick's is, I believe, therefore stronger in this respect. But Frankena's outline does make it possible to argue that Sidgwick's system could perhaps expand upon it. It also seems as though Frankena himself may well have agreed with such an endeavour. "It seems to me", he says, "that everyone who takes the moral point of view can agree that the ideal state of affairs is one in which everyone has the best life...in such state of affairs, it is clear that...both the principle of justice and the principle of beneficence will be fulfilled. If so, we can see that the two principles are in some sense ultimately consistent, and this seems to imply that increasing insight may enable us to know more and more how to solve the conflicts that trouble us..."¹²⁴. Frankena left to deontology a firm belief that while the concept of moral duty and obligation may be present at all times, so too is a certain importance of utility: So many deontological systems fail, Frankena says, on their simple failure to recognise this point.

0.8.b.iii. c) Robert Audi: Ethical Intuitionism

¹²² As will be seen in the next chapter, Frankena himself recognised that Sidgwick had pointed to a similar dissolution of the traditional boundaries between deontological and utilitarian understandings of morality.

¹²³ But what is particularly interesting in comparing these two systems is that Frankena, who is primarily considered to be a deontologist, begins from the principle of utility and arrives at duty, whereas Sidgwick, who is primarily considered to be a utilitarian, begins from the deontological principle of duty and only then derives the utilitarian principle.

¹²⁴ *Ethics*, p.73

Intuitionism has seen something of a revival in recent years, largely due to the work of Robert Audi. In his estimable work *The Good in the Right* (2004), Audi develops Rossian intuitionism further than Ross did himself, and to a point at which Audi integrates it with Kant's categorical imperative¹²⁵. The result is a strengthening of both Ross and Kant's positions, and a system that accounts for a theory of value as well as the duty based deontological position (that Audi simply refers to as ethical intuitionism) i.e. Audi's theory includes both the good and the right. If Audi is successful – which it is widely agreed that he is – then this is another form of reconciliation, and one that is unique in that it does not just look to (or result in) deontological and utilitarian/consequentialist views, but rather shifts the focus to a credible form of epistemological – and Kantian - intuitionism.

The first intuitionist to whom Audi refers as he is developing his understanding of intuitionism (and especially of the intuitionism that informed Ross), is actually Sidgwick himself. There are also similarities between Audi's appropriation of Kant for his intuitionist purposes and Sidgwick's, and in the fact that both philosophers incorporate a notion of the good into a largely Kantian framework. As part of my argument involves presenting Sidgwick's own intuitionism as being specifically comprised of both deontological and utilitarian properties however, the terminology of Rational Benevolence is likely to be different to Audi's ethical intuitionism. It is also on this basis that I argue that Rational Benevolence may be more directly relevant to the issue of the 'divide' between deontology and utilitarianism, if it is understood that that divide has been perpetuated by the unhelpful use of these mutually exclusive labels. But I am aware of the eminence of Audi's argument, and propose that further engagement with it in the future would be constructive for the arguments made here.

0.8.b.iii. d) Derek Parfit: Triple Theory

The last, most recent, and probably most important figure relevant to the field of reconciling deontological and utilitarian moral properties is Derek Parfit. Of all moral philosophers of the last thirty years, Parfit is among the most influential. His 1984

¹²⁵ Audi, *The Good in the Right*, (Princeton: Princeton University Press: 2004)

book *Reasons and Persons*¹²⁶ was regarded as something of a work of genius when it was released, and to this day remains hugely important for its rational arguments regarding wide moral problems such as population ethics, and responsibility towards future generations.

Parfit is primarily associated with a utilitarian position, but this is to greatly undersell the philosophic detail with which he approaches consequentialism in *Reasons and Persons*¹²⁷. His argument there is actually that consequentialism in all its present forms needs to be revised – and that common sense morality also requires this. Crucial to both of these revisions, Parfit argues, is a fundamental rationality – although he does profess in *Reasons and Persons* that he himself has failed to identify ‘Theory X’, which could solve the various problems he has identified in these ethical frameworks.

More recently however, in his long-awaited work *On What Matters* Parfit has presented a theory that goes a long way towards meeting the requirements of ‘Theory X’¹²⁸. This is an argument for the coalition of deontological, consequentialist and contractarian theories. As Samuel Scheffler says in the introduction to this two-volume colossus of a work, ‘Parfit aims to rechart the territory of moral philosophy’, by challenging the common assumption that there is a fundamental disagreement between consequentialists and Kantians, and by arguing for a ‘startling convergence’ among these positions instead¹²⁹.

There is not space in this thesis to engage with Parfit’s sophisticated arguments to the extent required to do them justice; to attempt to do so would be both philosophically and temporally too ambitious. However, Parfit’s theory is, ideologically at least, something close to the sort of reconciliation between deontological and utilitarian principles that I have been suggesting is needed, and an analysis of his position is crucial to any theory attempting to enter this field of reconciliatory deontological-utilitarian theories. Full engagement with Parfit’s theory

¹²⁶ New York: Oxford University Press: 1984

¹²⁷ For example, he introduces The Repugnant Conclusion (which he actually associates with Sidgwick), which denies total-utilitarian standards on the basis that if it is better that there are more people living to have some degree of happiness rather than less people living (even if the latter group have more happiness), then we are led through a series of population increases/happiness decreases, until we reach society Z, in which there are hundreds of billions of people alive, but all with lives barely worth living (*Reasons and Persons*, 381-391). He brings a similarly destructive case against average happiness utilitarianism.

¹²⁸ *On What Matters* (New York: Oxford University Press: 2011).

¹²⁹ Scheffler, in, *On What Matters*, p. xx

would be a possible line of future work; here I give a rudimentary outline of Parfit's theory, using the summary of *On What Matters* that Parfit himself provides at the beginning of the volume, and compare it to the method through which Rational Benevolence is established.

The beginning of Parfit's theory is provided by Kant's argument for universal laws. Parfit then works through several formulations of a Kantian basic maxim ('it is wrong to act on some maxim unless we could rationally will it to be true that everyone accepts this maxim, and acts upon it when they can'), and argues (the very non-utilitarian point) that Kant's formula, when revised in certain ways, can actually be made to be 'remarkably successful'¹³⁰. Parfit next adds the caveat of 'what if everyone did that?', and revises Kant's formula further to ensure an element of impartiality. This leads to the inclusion of the next major theory – contractualism – which works from a rational agreement formula based on Thomas Scanlon's theory that 'Everyone ought to follow the principles that no-one could reasonably reject'. This version of contractualism, Parfit argues, coincides with the Kantian formula, in which we should ask which principles each person would rationally choose, if they had the power to choose what everyone would accept. Lastly, Parfit addresses consequentialist theories, in which he defines the term 'best' as referring to the impartial-reason-implying sense that he has been developing. Then, with great simplicity, Parfit lays out how the revised contractualism formulation of the Kantian maxim admits of a form of rule consequentialism – "everyone ought to follow the principles whose universal acceptance would make things go best". That is, the principles that everyone would rationally choose are precisely those that inform this formulation of rule consequentialism. Thus, the result is a Kantian Rule Consequentialism¹³¹.

Parfit refers to this Kantian/contractualist/consequentialist combination as 'Triple Theory'. This is a 'single complex higher level property under which all other wrong-making properties can be subsumed. If this theory succeeds, it would describe what these other properties have in common'¹³². It would also, Parfit says, pose a strong challenge to the problem of moral disagreements. Where there have been those apparently deep disagreements between Kantians, contractualists and

¹³⁰ *On What Matters*, p.16

¹³¹ *Ibid*, pp.23-25

¹³² *Ibid*.

consequentialists, this is wrong. They have been, as Parfit says in a striking image, ‘climbing the same mountain on different sides’.

The similarity between what Triple Theory offers and what I am claiming Rational Benevolence is offering is obvious; both are reconciliatory theories that aim to close the gap between traditional deontological and utilitarian moral approaches, both claim that the disparity between them was erroneous in the first place – and both use a Kantian maxim *and* a principle that invokes impartiality¹³³. There are, however, two

¹³³ Brad Hooker disputes the strength of Parfit’s impartiality factor, on the basis that it is not sufficient to simply equate impartiality with agent-neutrality, as impartiality does not necessarily entail equal concern for everyone (and also that agent-neutral concerns can be ‘silly’ ones, that are not attractive morally) (Hooker, ‘Must Kantian Contractualism and Rule-Consequentialism Converge?’, in, *Oxford Studies in Normative Ethics*, Vol.4 (2014) 34-52). This leads to a situation in which the agent-neutral selection of rules within the contractualism that Parfit uses has ‘no justificatory force’ (p.48). Also, it is clear that attaching impartiality to agent-neutral evaluations will not necessarily produce convergence on rule selection, as agent-neutral concerns can still be different (p.49) Because of these differences in agent-neutral assessment, this may well lead to a disagreement about which rules everyone has the most agent-neutral reasons to accept (p.50). Ultimately, this would cause a divergence between Kantian contractualism and rule consequentialism, because one set of agent-neutral/impartial reasons will be guiding the view of one, and a different set the view of the other. Equally, it is possible for two Kantian contractualists to hold different sets of agent-neutral reasons, and therefore even these may disagree. (There is a parallel between this point of Hooker’s, and Sidgwick’s point about the flaw of the Kantian maxim, which is that it *still* allows for subjectivity, despite aiming at pure objectivity. See chapter 4 of this thesis). Unless the same impartial reasons drive both the contractualism and the rule consequentialism, then the two will not be able to converge. Impartiality then, according to Hooker, must be more than mere agent-neutrality. However, Hooker points out that provided agents take into account the right agent-neutral considerations, then this does result in convergence, and this *is* a remarkable conclusion for Parfit to have drawn (the same will apply even if the two positions only take into account the *same* considerations – they do not necessarily have to be the right ones). But for either of these things to happen, there must be agreement on which agent-neutral considerations are paramount (p.50). All of this leads Hooker to suggest that this sort of coherence might be best represented by a form of reflective equilibrium (pp.50-51).

As I have outlined above, Sidgwick’s theory of Rational Benevolence also includes an agent-neutral/impartiality aspect, firstly in the form of the self-evident principle of Prudence, which establishes temporal neutrality, and then Benevolence itself, which establishes agent neutrality. It is possible, therefore, that this part of Rational Benevolence would be subject to the same queries by Hooker as those he brought against Parfit. This seems especially likely given my argument that Rational Benevolence does something similar to Triple Theory by creating a convergence between a Kantian approach to morality, and a wider utilitarian one. My line of defence against this is to suggest that the utilitarianism of Rational Benevolence *does*, specifically and *only*, take into account the *same* agent-neutral concerns as Kantianism. I believe that Sidgwick himself expressly recognised the problem that Kantianism alone could not produce an adequate system for guidance, precisely because it did not seem to eradicate the problem of subjective moral beliefs, and that by rigorously testing what *could* be said of the Kantian maxim, Universal Benevolence is the only morally *rational* result. I theorise therefore that Rational Benevolence may avoid the problem of simply equating agent-neutrality with impartiality, because the particular impartiality in Rational Benevolence and its combination with Sidgwick’s negative version of the Kantian maxim produces only *one* ultimately reasonable agent-neutral concern. That is, Kantianism and consequentialism are, in Sidgwick’s theory, made to have the *same* agent-neutral concerns – and this is one way in which Hooker thinks Parfit has a valid point. Sidgwick’s theory is probably simpler than Parfit’s, but this may actually be to its advantage. In terms of whether this amounts to a form of reflective equilibrium, it will be seen (in chapter 1) that it is difficult to conclude whether or not Rational Benevolence is reflective equilibrium, because it is difficult to conclude whether Sidgwick is predominantly a coherentist, or a foundationalist.

points on which some distance appears between Triple Theory and Rational Benevolence. The first is that Sidgwick's theory is more traditionally philosophical than Parfit's, by which I mean it is more meta-ethical in an abstract way. As important as Parfit's Triple Theory is to deontological/utilitarian reconciliatory theories, I do not believe that it reconciles the moral properties in the same way as Sidgwick's argument for the truly mutually dependent nature of the relationship between deontological and utilitarian properties does. Whereas Parfit's argument is highly logical and formulaic, Sidgwick offers the basic, metaphysical means of understanding non-consequentialist duty to be inextricably attached to utilitarian considerations, which will be particularly effective if it can be shown that Rational Benevolence conforms to a synthesis in the way I have described it above. In that synthesis, the two rational moral values naturally both qualify and limit each other, which is how they are made to be irrefutable components of one coherent whole. This, as I have argued above, is a highly effective way of holding deontological and utilitarian principles together in such a way that they are in dynamic relationship rather than contention, which would be most helpful for the problem of moral ambivalence that I am trying to address. Parfit does also use the language of synthesis, but Triple Theory is not a synthesis of the kind I propose to draw out of Sidgwick's Rational Benevolence.

The second area in which Rational Benevolence presents something slightly different to Triple Theory is that of normative value. Rational Benevolence offers a potential method for being able to identify 'rationally benevolent' actions – actions that can be seen to include the values (and thus the limits) of both rational deontological and utilitarian properties and are thus more robustly morally defensible in contexts such as the Parental Predicament. Triple Theory is highly sophisticated, but it is unclear how this single, complex, higher level property is to apply at a practical level.

On these two counts, I believe that Rational Benevolence might still be able to contribute something of value to understanding how deontological and utilitarian properties can be united, even in the face of Parfit's accomplished theory.

0.8.c Other Considerations on Rational Benevolence as a Moral Theory

There are two remaining possible objections that could be brought against my argument for Sidgwick's theory of Rational Benevolence. The first is - what of other moral theories? The most obvious of these that is omitted from the above discussion is virtue ethics. It is true that virtue ethics arose largely in specific response to modern moral philosophy's preoccupation with moral obligation as it was being presented by deontological or utilitarian principles, and it is certainly one of the only positions (not including religious ones) that can genuinely claim to be approaching morality in a way different to those two major schools¹³⁴. It may even provide another possible means of challenging the traditional divide. However, where deontological and utilitarian normative theories are concerned with right action, the central questions of virtue ethics are those such as 'how should I live?', 'what is the good life?' and 'what are the correct personal and societal values?', and this does not often involve asking what it is right or wrong to actually *do* in a given situation. This question will be particularly pertinent in highly sensitive situations which are unexpected, and which are constrained by time in some way. The Parental Predicament, for example, is likely to involve one or both of those factors. In any case, I argue that normative questions are always going to be asked of *any* ethical situation in some capacity, and the answers to them are always going to involve recourse to the same options: Do we obey absolute principles? Or do we act in whatever way will secure the best outcome in the circumstances? I suggest that virtue ethics has very little constructive to say to these questions, and therefore to the problem of moral ambivalence, other than that cultivating the right character would enable an individual to act rightly in any situation, rather than just the specific situation in which they are in at the time – and I cannot see that this approach is adequate for providing a real guide for action in what is often the unanticipated here and now, or for offering better protection to parents within the framework of autonomy.

The last theoretical consideration of Rational Benevolence however must be this: Is Rational Benevolence, and Sidgwick's moral theory, ultimately just utilitarianism? This is, of course, the tide of prevailing opinion, and there are likely to be good grounds for such volumes. Sidgwick did at times align himself with a utilitarian

¹³⁴ See G.E.M. Anscombe, 'Modern Moral Philosophy', in, *Philosophy*, Vol.33, No.124 (1958) 1-19; Bernard Williams, *Ethics and the Limits of Philosophy* (London: Fontana: 1985); Philippa Foot, *Virtues and Vices* (Oxford: Blackwell: 1978); Alasdair MacIntyre, *After Virtue* (London: Duckworth: 1985)

position, although it is widely recognised that this was a highly sophisticated form of utilitarianism¹³⁵. But I agree with Schneewind that it does not necessarily follow that *ME* is itself a defence of utilitarianism. Sidgwick is clear about this from the outset, and it is my view that the results of the investigations in *ME* reveal to Sidgwick that utilitarianism is absolutely dependent on other self-evident principles, which he lifts directly from the deontological tradition, as I have outlined above and will argue in full in the thesis. I argue therefore that Rational Benevolence itself is *not* simply utilitarianism, and this is the case I will be making here.

On the other hand, it may be that Sidgwick *does* offer a form of utilitarianism, and that this might in itself be a perfectly acceptable version of the synthesis in Rational Benevolence. Brad Hooker argues that rule utilitarianism, and a form of reflective equilibrium therein, is an effective means of accounting for and unifying our various moral convictions. Hooker also argues that Sidgwick himself *could* be interpreted to be a rule utilitarian¹³⁶. Following Hooker then, it might be said that Sidgwick is indeed a type of rule utilitarian, and that this is actually to his great advantage. This may well also be another way of arguing that Sidgwick achieved the sort of unification between deontological and utilitarian moral inclinations for which I am arguing via Rational Benevolence. But to label Sidgwick as a rule utilitarian does not, on my view, pay adequately specific attention to the crucial deontological aspect of his work that makes his theory so distinctive, and so effectively overcomes ‘the antithesis’ between deontology and utilitarianism. I agree with Hooker that our primary moral intuitions need to be coalesced, and I believe this was also Sidgwick’s position - I have presented the case above that Sidgwick’s own epistemic intuitionism is itself

¹³⁵ See Sidgwick, ‘Professor Calderwood on Intuitionism in Morals’, p.564. Bart Schultz has also said recently that Sidgwick ‘set out with unparalleled clarity and force the utilitarian view that maximising happiness...was the ultimate normative standard’ (Schultz, *A more Reasonable Ghost*, <https://roundedglobe.com/books/34a3e7ff-778f-48d5-bca0-ed4e10132715/A%20More%20Reasonable%20Ghost:%20Further%20Reflections%20on%20Henry%20Sidgwick%20and%20the%20Irrationality%20of%20the%20Universe/> (no date; accessed 20/6/2016) p.2

I maintain however that even if Sidgwick is ultimately called a utilitarian, it appears evident that he has appropriated Kant to his cause to such a great extent that he has incorporated Kant’s ideas into his own view - and this does suggest that Sidgwick’s thought is at least sympathetic to deontological ideas.

¹³⁶ See Hooker, ‘Rule Consequentialism’ (Stanford), and ‘Sidgwick and Common Sense Morality’ pp.355-359. I have italicised ‘could’ in this sentence, however, because this does require some qualification. Overall, Hooker tends most often towards the argument that Sidgwick was an act utilitarian (see also *Ideal Code* pp.5-6, 99). But Hooker does believe that Sidgwick’s emphasis on assimilating common sense morality can accord with a rule-utilitarian approach (*Ideal Code*, pp.114, 116). Hooker also points out that in any case, rule-utilitarianism would have offered Sidgwick a better method for doing this than the act form (‘Sidgwick and Common Sense Morality’, p.360).

comprised of both deontological and utilitarian principles (as I will argue more extensively in chapter 2). But this is precisely why I dispute Sidgwick's utilitarian label. It is *only* because Sidgwick's theory of Rational Benevolence is a synthesis between equally important and mutually informing deontological and utilitarian concepts that it can be so directly applicable to the problems outlined at the start of this thesis. Because Rational Benevolence embodies this balanced meeting between the two moral properties, it accounts for the dual reasonableness that I have theorised to be the cause of moral ambivalence, and legitimises the rationality of both positions. And it is my argument that this elevates Sidgwick away from a merely utilitarian label.

Lastly, even if it is admitted that Sidgwick is not a utilitarian, there are a number of ways in which it might be argued that Rational Benevolence is not a 'synthesis' between deontological and utilitarian principles¹³⁷. It might be stated that the relationship is not truly mutually informing – that it is weighted in favour of one moral property or the other, or that it is not necessary that the deontological and utilitarian aspects both retain those qualities, *and* depend on each other. But it will be my aim in this thesis to build a case that the relationship between utilitarianism and deontology in Rational Benevolence *is* a synthesis of that nature. I also emphasise in response to those possible objections that Sidgwick himself used the language of synthesis, and that he recognised the importance of attempts to bring together the differing but apparently valid moral inclinations that make up the human moral experience. It was Sidgwick's hope that he could aid towards the construction of such a system, and it is my hope that this thesis might go some of the way towards constructing that system itself.

0.9. Summary of Aims, and Structure of the Thesis

¹³⁷ For example, David Phillips does not believe that Sidgwick achieved the reconciliation that he thought he had, on the basis that there is a fundamental problem with Sidgwick's argument for his own self-evident principles, from which he constructs his moral theory (Sidgwickian Ethics, pp.95-111). Thomas Hurka agrees with Phillips (Hurka, 'Sidgwick on Consequentialism and Deontology', in, *Utilitas*, Vol. 26, No.2 (2014) 129-152, p. 147-151), Both of their positions are considered in the literature review.

It is my principal aim in this thesis to show that Sidgwick's theory of Rational Benevolence provides a synthesis, understood as an interdependent and mutually informing relationship between utilitarian and deontological moral properties, that I have argued is required for a full understanding of the human moral experience. Specifically, this will involve first demonstrating how Sidgwick argues for the rationality of both deontological and utilitarian principles of morality, and then how this rationality unites them to create his moral theory of Universal Benevolence, which is best represented by the term 'Rational Benevolence'. I theorise that in this way, Rational Benevolence could offer a reconciliatory theory that challenges the traditional divide between deontological and utilitarian moral principles as it exists at the philosophical level, and also a means of addressing moral ambivalence in such a way that more than one course of action can be morally justified in ethically difficult situations.

There are three main stages to constructing this argument, which are laid out in the following chapter outline.

Chapter 1: Literature Review: Other Readings of Henry Sidgwick

The overall purpose of this chapter is twofold. Firstly, it aims to create a space in which to argue that there is a recognisably deontological aspect of Sidgwick's work that has largely been missed, due to the fact that he has always been read as a utilitarian. Secondly, it argues that Sidgwick's reconciliation between deontological and utilitarian moral properties has also been overlooked. The chapter begins from Sidgwick's time, and moves through the most notable theorists of the 20th century who have categorised Sidgwick as a utilitarian. The next part of the chapter focuses on the resurgence of interest in Sidgwick that has taken place in the last two decades, and specifically where Sidgwick has appeared in a significant way in the work of prominent theorists, but without having been utilised for his reconciliatory theory. The last part of the chapter examines recent theorists who have acknowledged that Sidgwick's work offers *some* kind of reconciliation, and argues that the particular synthesis of Rational Benevolence has not been recognised.

Chapter 2: Sidgwick's Account of Deontology

Here the case is made for there being a recognisably deontological influence in Sidgwick's work. I argue here that by 'Intuitionism' (with a capital 'I') Sidgwick is referring to a recognisably deontological method of ethics, understood as representing Reason-based non-consequentialism, and that this traditionally Kantian form of duty is actually found to contribute significantly to his own understanding of moral obligation. This relies almost exclusively on textual analysis, referring to *ME*, the personal document, and other articles and papers by Sidgwick. The second half of the chapter is concerned with examining the value that Sidgwick ascribes to deontological-based dogmatic intuitionism (or common sense morality) as a method of ethics that appears in the moral process of ordinary individuals, and assessing the place of this system of ethics in Sidgwick's overall moral theory. I aim here to show that despite showing most principles of dogmatic intuitionism to not actually be self-evident, Sidgwick does not himself entirely discount the method, and that he derives the principle of Universal Benevolence as an ultimately rational moral obligation from the self-evident, non-consequentialist bases of Justice and Prudence. Through this I also argue that Sidgwick's 'Intuitionism' actually informs his own epistemic 'intuitionism'. The chapter concludes that Sidgwick affords far greater weight and value to non-consequentialist moral properties than is commonly thought - but that Sidgwick does not consider deontological principles to be by themselves sufficient for constructing a wholly comprehensive or accurate account of morality.

Chapter 3: Sidgwick's Account of Utilitarianism

This chapter begins from the acknowledgement that Sidgwick is a utilitarian in some sense, and that this utilitarianism is of course essential to the proposed synthesis. I do however argue in this chapter that Sidgwick's discussions of utilitarianism are often less 'typical' than might be expected of classical or standard utilitarianism. Four defining features are identified in order to make this argument, that also guide this stage of the investigation. These are; that Sidgwick was not intending to write a defence of utilitarianism; that his predominantly Kantian understanding of rightness calls into question the status of means to ends; that Sidgwick recognises the 'defects and difficulties' of the utilitarian position; and that Sidgwick's own version of the utilitarian doctrine is a stringent one. I then argue that the last stage of his discussion on utilitarianism, in which Sidgwick aims to give the 'positive' proof for utilitarianism, can only be fully understood in light of the 'negative' relation between utilitarianism

and common sense morality given in Book III. That ‘negative’ relation (which I argue is not actually a negative relation at all, but another form of the positive relation) is a philosophical connection between the two moral positions that accounts for Sidgwick’s statement in Book III that ‘the truth lies between these conclusions’, and the sister passage in Book IV that states a line of argument is needed that does allow the validity to some extent of common maxims. This situation thus provides the starting point for the following chapter, which will argue for that philosophical connection – and the relationship between deontological and utilitarian principles in Sidgwick’s work.

Chapter 4: ‘Rational Benevolence’: Sidgwick’s Synthesis

In this chapter, all the claims made throughout the preceding stages of the thesis are brought together to fully form the arguments that Sidgwick perceived there to be a necessary and inherent relationship between deontological and utilitarian principles of morality, and that the nature of the relationship he creates conforms to the definition of ‘synthesis’ that I have given. The main discussion is taken up by a comprehensive analysis of *ME*’s Book III, in which Sidgwick uncovers the essential link between the fundamental, self-evident and non-consequentialist principles of Justice and Prudence, and the subsequently derived self-evidence of utilitarian Universal Benevolence. This connection, I will argue, is based on the inherent rationality of each type of moral property. It is only the rationality of the self-evident, non-consequentialist principles of Justice and Prudence that establish Universal Benevolence as ultimately rational; it is Universal Benevolence as the ultimately rational end of moral action that confirms and contextualises the place of Justice and Prudence. All must retain their original respective non-consequentialist or utilitarian properties in order to be mutually informing in this way – but all also depend on each other for full realisation and validation. In this way, the rightness of absolute-principle based and end-based moral properties are inextricably linked within a coherent whole that is most suitably referred to as ‘Rational Benevolence’. The last section of the chapter then turns to the PD, where Sidgwick outlines clearly the personal process through which he came to believe that there is no opposition between deontological and utilitarian principles – and in which he plainly states that neither the deontological position in Intuitionism, nor utilitarianism, are morally complete without reference to the other.

Conclusion

The results of chapter 4's argument for the synthesis embodied by Rational Benevolence are here examined in light of the problem of moral ambivalence that was outlined at the start of this introduction. Applying the findings of Rational Benevolence to the context of the Parental Predicament, it will be argued here that Sidgwick's synthesis between two different but equally rational moral properties directly reflects the dual reasonableness that causes moral ambivalence, and thus explains moral ambivalence/moral conflicts more effectively than theories that attempt to argue exclusively for either absolute-principle based or outcome-based approaches. That is, I argue here that Rational Benevolence provides a means of understanding both courses of action in situations of grave moral ambivalence to be, in a sense, 'right'. I then follow up this explanative function, and argue that Rational Benevolence has two levels of 'descriptive value'. These are a) the superficial, in which Rational Benevolence is simply explaining more accurately and more realistically the moral conflicts that we are seeing, and b) the substantive, in which that explanation itself specifically points to the dual reasonableness that could provide the basis for Rational Benevolence as a system of normative ethics. I then outline how Rational Benevolence might be normatively applied, including an examination of exactly what is meant by 'rationally benevolent actions', and whether or not a universalizability factor could be established. A full investigation into Rational Benevolence as a theory of normative ethics is beyond the scope of this thesis, and areas for possible future work are identified in light of this. But overall I suggest that the synthesis between deontological and utilitarian properties as it appears at the theoretical level could transfer directly onto moral actions at the practical level in such a way that each position both validates and limits the other to an appropriate degree, as it was said above is required by the dual reasonableness of moral ambivalence.

CHAPTER 1: LITERATURE REVIEW

OTHER READINGS AND INTERPRETATIONS OF SIDGWICK

This chapter will examine the place of Henry Sidgwick's work within moral philosophy. It is in part a history of ideas, in that it tracks where and why Sidgwick's work has been relevant in the century that followed him, and in part a critical engagement with the work of more recent theorists for whom Sidgwick's ethical theory has been a significant influence. The aim is to ascertain how Sidgwick's work has been predominantly interpreted, and to construct a place for my particular reading of Sidgwick that I develop in this thesis, which consists of these main points: That Sidgwick's moral theory is not exclusively utilitarian, that there is a recognisably deontological aspect to his work that contributes vitally to his own theory of moral obligation, and that his theory of Rational Benevolence offers a synthesis between deontological and utilitarian moral properties that could contribute significantly to overcoming the traditional divide between those ethical systems.

The theorists included in this chapter range from Sidgwick's contemporaries to theorists writing today, and the chapter follows a mainly chronological structure to support this. But the chapter is also designed to tell a certain story, about the way in which Sidgwick's work has been passed down through history. The story's main theme is what has been *missed* about Sidgwick's work, and as such it is, I believe, an original one. The narrative moves through stages that correspond with the main ideas in my reading of Sidgwick. As outlined in the thesis introduction, Sidgwick has historically and (almost) universally been considered to be a classical utilitarian (of the Benthamite/Mill tradition), despite his work having been utilised by some distinctly non-utilitarian positions. In the 20th Century sections of the chapter therefore, attention will be drawn both to the prevalence of Sidgwick's utilitarian status and to the places where he was also momentarily relevant to deontological theories also, with the argument being made that the significance of that flexibility was overlooked. When it comes to the recent resurgence in interest in Sidgwick, and in the movement towards making him actively relevant to contemporary philosophy, it will be seen that his utilitarian reputation has endured. Whereas there has been some recent recognition of the fact that *ME* should perhaps not be read simply as a defence of utilitarianism, and even a degree of acknowledgement that Sidgwick might offer some kind of reconciliation between deontological and utilitarian

principles, neither of these aspects, still, have been fully realised. This account of Sidgwick's story will end therefore on the points that Rational Benevolence has not been recognised as the source of reconciliation in Sidgwick's work, that it has not been read as a *synthesis* between deontological and utilitarian principles – and the idea that there is now space in the field for these arguments to be made. It is to this part of Sidgwick's story that I intend this present thesis to be directly relevant.

In terms of the most recent literature, the focus is mainly on theorists who have addressed in some way the idea of Sidgwick's applicability to both deontological and utilitarian fields - but there are other writers included here, who although they do not directly discuss reconciliation have nonetheless presented readings of Sidgwick that have shaped his place in moral philosophy, and contributed helpfully to my reading of him here.

1.1. *The Methods of Ethics*: Reception and the Early Years

The first part in telling the story of how Sidgwick's work has been passed on through the years must of course take account of how *ME* was originally received, upon publication and in the years in which it was still relatively new. There are some important details here, in Sidgwick's discourse with his peers and critics, and in the criticisms themselves, that begin to build the path that *ME* would be set upon for much of its future. In these sections I take account of the reception of *ME*, and of where it appeared in the work of theorists whose ideas were to contribute significantly to the development of moral philosophy in the 20th Century.

1.1.a. Initial Reactions

By the time *ME* was published in 1874, Sidgwick had been working on it for over a decade¹³⁸. Sidgwick was careful to distinguish that by the 'The Methods' of the title he did not mean *a* or *the* definitive method, but rather impartial observation and investigation into morality as it appeared and operated in ordinary moral thought. *ME* was meant to be a 'technical work'¹³⁹, which would introduce clarity into an otherwise highly confused and diverse field. Sidgwick intended it to be the most comprehensive

¹³⁸ As mentioned in the thesis introduction, Sidgwick alludes this work as early as 1862, in the letter to H.G. Dakyns. He also told several other people over the next few years that he was developing a "great work" (Schultz, *Eye of the Universe*, p. 141).

¹³⁹ See Schneewind, *Sidgwickian Ethics*, p.191

and thorough inquiry into the nature of the human moral experience yet written – and it was. Sidgwick’s systematic approach to moral thought, and his diligent attention to detail, bore *ME* upon its publication to the heights of instant recognition.

The initial discussion and reviews of *ME* - though vast – are varied. Much of the literature is positive regarding the volume’s scope and depth. H. Calderwood presented a fierce challenge to Sidgwick’s version of intuitionism, but elsewhere states his ‘admiration of the ability manifest everywhere throughout the book’¹⁴⁰. A. Bain, in his long scrutiny of *ME*, praises Sidgwick for ‘not blinking any difficulties’, that arise during the investigation of each method (hedonism, in particular), and despite criticising some of the later aspects of *ME*, also accredits it with being one of the best volumes of its generation¹⁴¹. This sort of statement was to become a running theme in *ME*’s lasting reputation¹⁴². In terms of Sidgwick’s own contribution to moral philosophy (as opposed to *ME*’s value as a sort of encyclopaedia of the various ethical approaches) however, *ME* divided opinion. F.H. Bradley took Sidgwick to task on the controversial issue of hedonism¹⁴³. For some, Sidgwick’s critical and non-constructive approach to philosophy rendered him fairly useless in terms of true meta-ethics¹⁴⁴. In a review of Sidgwick’s later work *Philosophy: Its Scope and Relations*, A.E. Taylor said of *ME* that the ‘cardinal defect’ of the volume was its failure to fundamentally grasp some philosophical problems¹⁴⁵. Even C.D. Broad was later to venture that some of *ME*’s claims were ‘extraordinarily trivial’¹⁴⁶. It

¹⁴⁰ Calderwood, H. ‘Mr Sidgwick on Intuitionism’, in, *Mind*, Vol.1, No.2 (1876) 197-206, p. 206

¹⁴¹ A. Bain, ‘Mr Sidgwick’s Methods of Ethics’, in, *Mind*, Vol.1, No.2 (1876) 179-197, p. 185

¹⁴² Hastings Rashdall stated *ME* to be an “instant philosophical classic” in 1885 (‘Professor Sidgwick’s Utilitarianism’, in, *Mind*, Vol.10, No.38 (1885) 200-226, p.201), C.D. Broad famously declared it to be ‘the best treatise on moral theory that has ever been written’ in 1930 (Broad, *Five Types*, p.143), and John Rawls called it ‘the first truly academic work in moral theory, modern in both method and spirit’ (Rawls, Preface to Sidgwick’s *The Methods of Ethics*, 7th Edition (Indianapolis: Hackett Publishing: 1980), p.v). See also Henry Barker, ‘A Recent Criticism of Henry Sidgwick’s Methods of Ethics’, in *Philosophical Review*, Vol.11, No.6 (1902) 617-603; William Pepperell Montague, *The Ways of Knowing, or The Methods of Philosophy* (London: Compton Printing Works Ltd: 1925) p.5, and Brand Blanshard, *Reason and Goodness* (London: George, Allen, and Unwin: 1961), p.90 - all of whom are united on the basis of great regard for Sidgwick’s treatment of ethics, even if they are divided about the results of his arguments.

¹⁴³ F.H. Bradley, ‘Mr Sidgwick on Ethical Studies’, in, *Mind*, Vol.2, No.122 (1877)

¹⁴⁴ Schultz states that much of the controversy surrounding *ME* when it was first published arose because Sidgwick tended to keep philosophy ‘at arm’s length’ (Schultz, *Eye of the Universe*, p.142). This is connected, perhaps, to Sidgwick’s preferring to always keep arguments open, and to the fact that he wasn’t aiming to give an ethical theory, but rather an exploration.

¹⁴⁵ A.E. Taylor, ‘Review: *Philosophy: Its Scope and Relations: Henry Sidgwick*’, in, *Ethics*, Vol.13, No.3 (1903) 377-, p.378. Taylor was still ready, however, to admit to Sidgwick’s ‘genius’, that showed in all he had written as a critical philosopher.

¹⁴⁶ Despite his initial glowing review, it is also Broad who draws attention to Sidgwick’s ‘grave defects as a writer’ (*Five Types*, p.144). These are almost universally accepted as over-complexity, an

was vaguely recognised that, despite apparently following in the tradition of Bentham and Mill in some respects, Sidgwick was aware of the difficulties of the utilitarian doctrine, and was clearly prepared to address them in a far more credible way than either Bentham or Mill had attempted. But so much of all of this was overshadowed by that now infamous Sidgwickian problem, the Dualism of Practical Reason (here after 'DoPR') - which represented Sidgwick's inability to reconcile reasonable self-interest with duty, and which Sidgwick himself loudly lamented¹⁴⁷. Sidgwick was persuaded to alter the last chapter of *ME* in order to reduce the emphasis on this problem¹⁴⁸. But despite these revisions, the DoPR was to remain among the most persistently reviewed of all aspects of Sidgwick's work, and arguably overshadowed some of *ME*'s more refined achievements¹⁴⁹.

Overall, *The Methods* obtained a place of immediate importance upon its publication, and retained a significant degree of philosophical merit thereafter. Interestingly, not much of the literature contemporary to Sidgwick makes much of *ME*

impenetrable density of content – and an unfortunate tendency to dullness. Broad also believed Sidgwick, on some points, to have failed to actually produce any truly substantial answers to some of the more pressing questions. But Broad's more generous assessment of Sidgwick also draws attention to another reason why *ME* earned the status that it did. 'His capacity for seeing all sides of a question and estimating their relative importance was unrivalled....and he never allowed the natural desire to make up one's mind on important questions to hurry him into a decision....Those who, like the present writer, never had the privilege of meeting Sidgwick can infer from his writings, and still more from the characteristic philosophic merits of such pupils of his as McTaggart and Moore, how acute and painstaking a thinker and how inspiring a teacher he must have been' (Broad, *Five Types*, p.144). Broad's assumption is proven by other more personal accounts of Sidgwick to be absolutely sound.

¹⁴⁷ Others also recognised the extent of the problem (see Lesley Stephens' review of *ME*, 'Sidgwick's Methods of Ethics' in *Fraser's Magazine* (1875) n.s.11, for example), although none were harsher critics of this problem than Sidgwick himself.

¹⁴⁸ During the section in the preface to the second edition in which Sidgwick discusses the amendments he had made to the last chapters of *ME*, he reminds his readers of Butler's influence, but also emphasises where he differs from Butler. According to Sidgwick, Butler had never seriously asked the question 'what among the common precepts of our common conscience do we really see to be ultimately reasonable?' whereas Sidgwick himself *had* asked this question. What is most important about this is that it leads to Sidgwick stating – plainly and without qualification – that the answer he found 'supplied the rational basis that I had long perceived to be wanting to the Utilitarianism of Bentham....and thus enabled me to transcend the commonly received antithesis between Intuitionists and Utilitarianism'. That is, there is very strong evidence here to support the argument that the DoPR not only need not stand in the way of drawing out of Sidgwick's work a synthesis between deontological and utilitarian principles, but that Sidgwick's deeper investigation of into what can be called ultimately reasonable actually *yielded* such a result. This point will be resumed in subsequent chapters.

¹⁴⁹ See, as just a few examples of works that are predominantly concerned with the DoPR, W.R. Sorley, *A History of English Philosophy* (New York: G.P Putman's Sons: 1921); J.L. Mackie, 'Sidgwick's Pessimism', in, *Philosophical Quarterly*, Vol.26, No.105 (1976) 317-27; David O Brink, 'Sidgwick and the Rationale for Rational Egoism', in, Bart Schultz, *Essays on Henry Sidgwick* (Cambridge: Cambridge University Press: 1992) pp.199-240; W. Frankena, 'Sidgwick and the Dualism of Practical Reason', in *The Monist*, Vol.58 (1974) 449-467; Alison Hills, 'The Significance of the Dualism in Practical Reason', in, *Utilitas*, Vol.15, No.3 (2003) 315-329.

as a defence of Utilitarianism. This was to follow however, as the result of the aforementioned F.H. Bradley's issue with Sidgwick's hedonism, and Sidgwick's apparent incompatibility with another branch of moral philosophy that was developing at the time – Idealism.

1.1.b. Sidgwick's Utilitarianism and British Idealism

In the preface to *ME*'s second edition, Sidgwick responds to the criticisms from A. Bain and H Calderwood that were outlined above¹⁵⁰. But Sidgwick also references a third type of critic, who he doesn't mention by name. These critics have, Sidgwick says, despite Sidgwick's insistence in the preface to the first edition of *ME* that he was not intending to defend any one particular ethical theory more than another, 'overlooked or disregarded the account of the plan of my treatise...and [have] consequently supposed me to be writing as an assailant of two of the methods which I chiefly examine and a defender of the third' (*ME*: x). Chief of the critics to whom Sidgwick was referring here is undoubtedly F.H. Bradley.

Bradley was among the members of the first wave of British Idealism that began in the 1870s. Arguing a belief in an all-encompassing 'Absolute' reality, and concerned to emphasise the faculty of Reason as the means of accessing and understanding this Absolute, Idealism was to some extent a revival of the of the German idealism of Immanuel Kant and G.W. F. Hegel. More specific to the discussion here is that it was in part a reaction to the utilitarianism of David Hume, J.S. Mill – and Henry Sidgwick. In his 1877 pamphlet *Mr Sidgwick's Hedonism; An Examination of the Main Argument of The Method of Ethics*, Bradley took Sidgwick's apparent hedonism to be paradigmatic of that which the Idealists were rejecting, and mounted a sustained attack¹⁵¹.

¹⁵⁰ Sidgwick actually gave a very lengthy response to what he considered to be Calderwood's serious 'misapprehension' of his 'meaning and drift', in 'Professor Calderwood on Intuitionism in Morals', in, *Mind*, Vol.1, No.4 (1786) 563-566.

¹⁵¹ (Cambridge: Cambridge University Press: 1877). Schneewind points out that this was really the only protracted criticism encountered by *ME* in Sidgwick's lifetime. It was, however, fierce. Bradley had actually already criticised Sidgwick's account of reason in *Ethical Studies* (1876) In the evocative words of D.G. James, "Bradley trained on Sidgwick's book in a footnote to an essay in *Ethical Studies* a brilliant burst of machine gun fire; and then, in a brochure he put out in 1877, he brought up his full armoury and blew Sidgwick's position to pieces" (James, *Henry Sidgwick: Science and Faith in Victorian England*, (Oxford: Oxford University Press: 1970) p.33.

Bradley's main charge against Sidgwick was that his account of Reason was vague. (For others of this opinion, see E. Albee, *A History of English Utilitarianism* (Abingdon-on-Thames: Routledge: 1901) and James Seth ('The Ethical System of Henry Sidgwick') who also considered Sidgwick's rendering

Bradley's criticisms of hedonism were, and still are, considered to be seminal¹⁵². But if Bradley had introduced Idealism, it was T.H. Green – a colleague of Bradley's at Oxford - who made the most important developments to it. Despite having been great friends since both attending Rugby School, it was actually Green who provided a systemised Idealistic account of morality that could be thought to really compete with Sidgwick's. Heavily anti-utilitarian, on the basis that that doctrine lacked any metaphysical grounds, and concerned to show that human existence is not entirely governed by the laws of nature Green engaged extensively with ME in his *Prolegomena to Ethics*¹⁵³. Unlike Bradley however, Green held ME in great esteem. This was due to the crucial – and telling - fact that he did *not* consider Sidgwick's utilitarianism to be the standard classical doctrine. Green was, like Bradley, also concerned with Sidgwick's use of Reason - but where the original utilitarians such as Hume had believed Reason to be a faculty of means, rather than of ends, Green recognised Sidgwick's interpretation of utilitarianism to be highly Kantian, in that the function of Sidgwick's Reason was to yield ultimate *ends*¹⁵⁴. As Green describes it,

of Reason to be ambiguous. See Barker ('A Recent Criticism') for the completely opposite opinion). Bradley could not fathom whether Sidgwick thought particular moral judgements, or universal judgements, were the functions of Reason – and was also unclear on how Reason and desire were related in motivating an individual to act. Bradley's attempted definition of Sidgwick's Reason is "I think that Reason stands for the faculty which apprehends 'universal truth,' general rules', 'abstract moral notions', 'moral axioms'; and further that it includes as well the reasoning to particular cases from 'general moral axioms' whether these latter are 'universal moral intuitions' or not. The word may cover more, but I doubt it" ('Sidgwick's Hedonism', p.9). In Bradley's defence, that interpretation of Sidgwick's use of Reason is not entirely inaccurate; Sidgwick does often use an amalgamation of reasoning and intuitionism in his epistemology, and at other times separates them entirely. But Sidgwick was entirely capable of defending himself. The dialogue between the two philosophers, consisting mainly of reviews of each other's work, reveal Sidgwick at his quietly disparaging best. In his review of Bradley's *Ethical Studies*, Sidgwick described the book (and by extension Bradley himself) 'vehemently propagandist' (*Mind*, Vol.1 (1876) p.545). Of Bradley himself, Sidgwick said "penetrating criticism, especially in ethics, requires a patient effort of sympathy, which Mr Bradley has never learned to make, and a tranquillity of temper which he seems incapable of maintaining" ('Review of Ethical Studies', p.545)

¹⁵² See Stewart Candlish, 'The Truth About F.H. Bradley', in, *Mind*, Vol.98, No.391 (1989) 331-348, and Marcus Singer, 'Henry Sidgwick', 1992, p.73, in, Singer (ed.) *Sidgwick, Essays on Ethics and Method* (Oxford: Clarendon Press: 2000), p.xxxviii

¹⁵³ Green (Oxford: Clarendon Press: 1883). See Alasdair MacIntyre, *A Short History of Ethics* (Abingdon-on-Thames: Routledge & Kegan Paul Ltd: 1967) p.245.

¹⁵⁴ Green gave one of the most robust defences of Sidgwick's rejection of egoistic hedonism in favour of universal hedonism, drawing attention to the fact that it was Sidgwick's interpretation of Reason that had led him to that conclusion (Green, *Prolegomena*, pp. 409-10). Green believed most theories of the 'good' to be unavoidably circular, but allowed Sidgwick's theory to escape this accusation on the grounds that Sidgwick concluded that it is reasonable to seek as ultimate good a form of life that is reasonably desired. Whereas this may still appear circular, for Green this was Sidgwick agreeing with his own belief – highly Kantian and almost transcendental - that Reason gives its own end, and needs no further qualification.

Nor did Sidgwick's apparent alignment with the Idealist values of Green – and of Bradley – end there. It has since been proposed, by Dr C. J. Dewey perhaps most notably, that the Oxford Idealists

Sidgwick ‘accepted principles, as it would seem, so antagonistic to those of the philosophic utilitarians’, by having sought a ‘proof’ for the utilitarian principle which he only found to exist in Reason - but (inexplicably, to Green) ended ‘by accepting their conclusions’¹⁵⁵.

Green’s reading of Sidgwick was thus vital for suggesting that Sidgwick was only a very particular kind of utilitarian. There were sundry agreements with this idea:

Sidgwick’s friend James Bryce, for example, wrote of him – “the system of Utility takes in his hands a form so much more refined and delicate than was given to it by Bentham and James Mill, and is expounded with so many qualifications unknown to them, that it has become a very different thing.”¹⁵⁶ As noted in the introduction, F.H.

(Bradley and Green) actually shared with the ‘Cambridge Revisionists’ (of whom Henry Sidgwick was the main one) a common moral and social philosophy, on the basis that both groups rejected aspects of traditional utilitarianism (Dewey, ‘Cambridge Idealism’: Utilitarian Revisionists in Late Nineteenth-Century Cambridge’, in, *The Historical Journal*, Vol.17, No.1 (1974) 63-78. See also S. Collini, ‘Idealism and “Cambridge Idealism”’, in, *The Historical Journal*, Vol.18, No.1 (2009) 171-177, pp.171-172). Collini rightly points to several facts that seriously undermined Dewey’s claim, such as the vehemence of the debate between Sidgwick and Bradley in the pages of *Mind*, and the fact that they weren’t as unified on the rejection of hedonism as Dewey seems to assume that they are. But theories such as Dewey’s are important for the fact that if nothing else, they point to how easy it was to distinguish Sidgwick from the earlier utilitarians – to the extent to which writers such as staunch Idealist Green could actually draw parallels with Sidgwick’s views.

¹⁵⁵ Green, *Prolegomena*, p.411. Despite Green’s apparent support for Sidgwick on this point, Green’s ethics were a life-long battleground for Sidgwick, and the two had actually long been in dispute about the nature of ‘ultimate good’, which started after the publication of Sidgwick’s 1877 article ‘Hedonism and Ultimate Good’ (*Mind*, Vol.2, 27-38). Sidgwick and Green were also opposed on notions of pleasure, feeling, time as an abstract concept, and moral ideas, among other things. Before it became obvious in Green’s *Prolegomena* that Green was trying to recruit Sidgwick to a more transcendentalist position, the discourse had been long and convoluted. See Sidgwick, ‘Green’s Ethics’, in *Mind*, Vol.9, No.34 (1884) 169-187; Sidgwick, *Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays*, (London: Macmillan: 1905) 220-221; Sidgwick, *Review of Green and Grose’s Edition of Hume’s Treatise* (Academy, 30th May, 1874); Sidgwick (E.E. Constance Jones, (ed.), *Lectures on the Ethics of T.H.Green, Mr Herbert Spencer and J. Martineau* (London: Macmillan: 1902), pp.11-14, 15-22, 179.

In his review of Sidgwick’s *Lectures on the Ethics of T.H.Green*, James Seth states that although Green gave enormous weight to Sidgwick’s use of Reason, Green’s Idealism would forever be the only system of ethics on which Sidgwick never got a proper grasp (Seth, ‘Reviewed Work: Lectures on the Ethics of T.H. Green, Mr Herbert Spencer and J. Martineau’, in, *The Philosophical Review*, Vol.12, No.5 (1903) 548-553, p.550). Others however, such as D. Weinstein, look more favourably on the relationship between Green and Sidgwick, claiming that they both recognised that they shared “good-promoting” theories of morality, and that most of their arguments were actually due either to misunderstanding each other, or to each feeling that the argument of the other was vague to the point of incoherence (Weinstein, *Utilitarianism and the New Liberalism* (Cambridge: Cambridge University Press: 2011) pp. 53-55). Weinstein also draws attention to the point made in the main text above, that in terms of accepting the normative value of utilitarianism, Green moved closer to Sidgwick’s position than Sidgwick did to Green (*Utilitarianism*, p. 55).

¹⁵⁶ James Bryce, ‘Henry Sidgwick’, in, *Studies in Contemporary Biography* (London: Macmillan: 1903). It is also important to note the influence on Sidgwick of F.D. Maurice, an eminent figure among the Cambridge ‘Apostles’ (of whom Sidgwick was also a member) whose Christian faith led him to believe that much of life was undefinable. As Schultz remarks, Maurice was standing “behind both Mill and Sidgwick, as a powerful voice pleading the limits of utilitarianism” (Schultz, *Eye of the Universe*,

Hayward urged that ‘Sidgwick should not be classified as this or that....it is extremely difficult to classify him at all’¹⁵⁷. But the fact remains that in terms of Sidgwick’s moral position in his own day, whether in Bradley’s work or Green’s, the point of departure between Sidgwick and the Idealism of his time was that of hedonism. Even Alasdair MacIntyre, in his *A Short History of Ethics*, places Sidgwick ahead of Green and Bradley on the timeline, the implication being that Sidgwick was merely the last defender of utilitarianism in the face of rising Idealism, and that once Idealism took over, engagement with Sidgwick’s work was simply dropped¹⁵⁸. Of course this is chronologically correct in terms of published works, but the regular overlooking of Green’s rather more sympathetic reading of Sidgwick excludes the importance of that nuanced interpretation in favour of Bradley’s fundamental and more insistent claim that Sidgwick was a hedonist, and as such a utilitarian. Bradley’s attitude is indicative of the sort of engagement with *ME* that was to follow in the long term. D.G. James, writing on the status of Sidgwick’s work as it stood at the end of the classical era, speaks of Sidgwick’s ‘rehabilitation of utilitarianism’. ‘A utilitarian ‘science’ of behaviour’, James states, ‘with its show of cause, effect, quantities and measurements has been reaffirmed with the aid of a priori concepts, and has become, in Sidgwick’s hands, more catholic, respectable and traditional than it had ever had been before’¹⁵⁹. And where Sidgwick might not have disputed his utilitarian label per se, he would most likely have argued against the fact that this could be

p. 46). Sidgwick took note of Maurice’s careful approach to theological epistemology, and can be thought to have applied similar care in his own systematic treatment of each ethical position.

¹⁵⁷ See Hayward, ‘Mr Hayward’s Evaluation of Professor Sidgwick’s Methods of Ethics: A Reply’ in, *International Journal of Ethics*, Vol.11, No.3 (1902) 360-365. This is itself a reply to E.E. Constance Jones’ reply to Hayward’s original article ‘The True Significance of Sidgwick’s Ethics’ (in, *International Journal of Ethics*, Vol.11, No.2 (1901) 175-187). Hayward accuses Constance Jones - Sidgwick’s most devoted student - of assuming that Sidgwick’s departures from utilitarianism were simply ‘common sense’, when actually they were the result of a very specifically Kantian type influence. “I should rather regard this Rationalistic terminology as somewhat foreign to Hedonism. I do not think that Miss Jones will find, in Sidgwick’s Hedonistic predecessors, any such emphasis on Reason’ (‘A Reply’, p.361). Hayward thought that Sidgwick’s work was ‘unconsciously Kantian’, and that this was the result of Sidgwick’s emphasis on Reason. Schneewind has also claimed that Sidgwick’s view was importantly different to other utilitarians (*Sidgwickian Ethics*, pp. 201-21), and that there were some ‘markedly non-utilitarian features’ (see also Schneewind, ‘Sidgwick and the Cambridge Moralists’, in, *The Monist*, Vol. 58, No.3 (1974) 371-404, p.372). Schneewind also agrees with my position that Sidgwick’s philosophical intuitions (Justice, Prudence, and Universal Benevolence) are largely Kantian in function - ‘by requiring us to generalise whenever we assign a reason for an act or a desire’. A large part of Schneewind’s interpretation however is due to his contextualising of Sidgwick within the wider framework of Cambridge moralists of his day, and not exactly a direct comment on Sidgwick’s moral theory.

¹⁵⁸ MacIntyre, *Short History*, p.163

¹⁵⁹ James, *Henry Sidgwick*, pp.35-36

surmised from *ME*. Sidgwick himself summarises it perfectly; ‘Oh how I sympathise with Kant’, he once said, ‘with his passionate yearning for synthesis, and condemned by his reason to criticism’¹⁶⁰.

1.1.c. Sidgwick at the turn of the Twentieth Century: G.E. Moore and Bertrand Russell

Some of the influences that would be most formative for *ME*’s continuing reputation came from two of Sidgwick’s pupils – G.E. Moore, and Bertrand Russell. With these two figures, the idea would be continued that Sidgwick was a very particular kind of utilitarian.

By the time Moore and Russell were writing at the beginning of the 1900s, moral philosophy was moving rapidly away from the Idealism that had characterised it during Sidgwick’s era, and towards a more analytical and systematic form¹⁶¹. In Sidgwick’s academic sphere of Cambridge, Moore was at the head of this change¹⁶². Although Moore’s relationship with Sidgwick was complicated, his hugely influential *Principia Ethica* heavily cites *ME* on a point for which Moore was to be best known – his account of the good¹⁶³. For Moore, good is the fundamental ethical value, and as

¹⁶⁰ Letter of 1868.

¹⁶¹ Alasdair MacIntyre’s evocative description of moral philosophy’s arrival into the twentieth century is that it ‘opened on a quietly apocalyptic note’ (*Short History*, p.159). The sense was that moral philosophers had failed to answer the fundamental questions that had been asked so far – those of, ‘what kind of actions ought we to perform?’, and ‘what kind of things ought to exist for their own sake’. As Schultz describes it, Sidgwick was “cast as a kind of bookish, academicized remnant of this legacy, holding out against the wave of philosophical idealism that swept such figures as Green and F.H. Bradley into the forefront of British philosophy, until with the new century G.E. Moore and Bertrand Russell shifted the current, and contemporary analytical philosophy was launched” (Schultz, *Eye of the Universe*, p. 3).

¹⁶² Moore was – and remains – highly influential. So important was he in fact to the framing of moral philosophy in the twentieth century, that he has been credited with everything from singlehandedly changing ethical theory and ‘setting the trend’ for modern moral philosophy in general, (Schultz, *Eye of the Universe*, p. 5), to ‘creating the puzzles’ that everyone else would later try to solve (Joseph Margolis, ‘G.E. Moore and Intuitionism’, in, *Ethics*, Vol.87, No.1 (1976) 35-48). Moore had perceived that problem pointed out by MacIntyre, that important philosophical questions had not yet been answered, and theorised that this was because they had not yet been properly framed: His seminal work, *Principia Ethica* (Cambridge: University Press) published in 1903, was intended to do this. His writing style, which was fierce and direct, was naturally convincing; *Principia Ethica* was, according to Schultz, held up almost as an ‘object of worship’. (Schultz, *Eye of the Universe*, p. 5)

¹⁶³ Of all the people in the world not to like Sidgwick – and by all accounts not liking Sidgwick was a difficult thing to do – Moore, as his most famous pupil, is the most remarkable. But Moore was outspoken on the fact (Ann Dawes, *Henry Sidgwick* (Cambridge: Biograph: 2007) Moore wrote of his dislike of Sidgwick in his autobiography, and is described by Everett Hall as being possibly “the only person in the world who found his personality unattractive” (Hall, ‘The Philosophy of G.E. Moore’, in, *Philosophical Review*, Vol. 53, No.1 (1944) 62-68 p.62). Moore considered Sidgwick’s famous consideration for all viewpoints not to be a sign of great tolerance and patience, but of weakness, and lack of conviction. As for Sidgwick’s response to Moore, Sidgwick stated in a letter to Daykns that he

such cannot be defined via any describable, empirical properties: Moore's famous 'naturalistic fallacy' is his belief that it is an error to attempt to do so. Good is only identifiable via reference to 'moral intuitions', understood as self-evident propositions which cannot directly be either proven or disproven – a point on which he clearly, and self-confessedly, followed Sidgwick¹⁶⁴. Moore's positing of these intrinsic goods (or 'ideals', such as justice, freedom and knowledge) alongside a straightforward consequentialist account of the relationship between those goods and the 'right' actions that secure them, resulted in the first form of one of utilitarianism's most popular contemporary interpretations, 'ideal utilitarianism'¹⁶⁵. Most significantly for this current discussion, Moore's use of Sidgwick depended on his recognition of the Kantian influence in Sidgwick's work. Moore drew a direct parallel between Kant's argument that moral rules that pass the categorical imperative are ultimate definitions, and his own unanalysable account of the good, both of which also appear in similar guises in *ME*. Thus Sidgwick's precise and semi-Kantian rendering of both Intuitionism and utilitarianism was essential to Moore's own, and this was only possible on the basis that Sidgwick's non-naturalistic version was so different to that of the utilitarians that had gone before him. What is most striking about Moore's having identified this distinctive quality in Sidgwick's utilitarianism is his claim that even Sidgwick himself had not been aware of it. As Moore said in 1903, 'Sidgwick himself seems never to have been clearly aware of the immense importance of the difference which distinguishes his Intuitionism from the common doctrine'¹⁶⁶. Yet

hadn't even known Moore was publishing on ethics (Dakyns, 3rd Feb 1900), and died before Moore published *Principia Ethica* in 1903.

¹⁶⁴ Theories that argue that certain things are 'good' may well be correct, but they have not defined good per se – they have simply shifted the question of the property of goodness onto whatever it was they were using to define 'good'. According to Moore, Sidgwick had not committed the naturalistic fallacy, given that Sidgwick thought notions of 'right' and 'ought' were irreducible to definable terms. Schultz points out that many theorists think Sidgwick to have actually only produced an irreducible account of 'right', whereas Moore produced an irreducible account of 'good', which is an interesting interpretation given that both Moore and Sidgwick are held to be utilitarians.

¹⁶⁵ Although it is most commonly associated with Moore, the actual term 'ideal utilitarianism' did not appear until Hastings Rashdall used the phrase in his *The Theory of Good and Evil* (Oxford: Clarendon Press), which was published in 1907. Moore also argued that because it is always so difficult for us to determine by ourselves what will secure the best outcome, we would do best if we followed certain established rules. He is not far from Mill in this claim, but Moore would be retrospectively identified as an early 'rule consequentialist' more frequently than Mill, given his influence on modern utilitarianism in general.

¹⁶⁶ *Principia Ethica*, p.5. Moore goes on to explain that Sidgwick does not hold that actions that are held to be right, or 'duties', are incapable of proof or disproof, and thus does not follow the 'Intuitionists proper' as Moore called them. Instead, Moore considered himself to be united with Sidgwick in holding that only ethical intuitions as to what is ultimately good are *real* moral intuitions. Moore did not consider himself to be an intuitionist of the type that had preceded him – and he did not

even these words from the influential Moore did nothing to stop the tide of opinion that Sidgwick was a classical – and increasingly outdated – utilitarian.

The impact that Moore *did* have on the perception of Sidgwick's work is an important and somewhat diverse one. As Hurka describes, Sidgwick was actually often seen to be defending an 'unanalysable right', whereas Moore was defending an 'unanalysable good' but this had either not mattered, or been apparent, to Moore¹⁶⁷. Moore himself went some way to recognising this when he drew on Sidgwick's Kantian notion of irreducible notions, but ultimately his emphasis on the good contributed substantially to furthering the 'right' vs 'good' aspect of the deontological/utilitarian divide - and Sidgwick was heavily indicted in this. Sidgwick may have been a very particular type of utilitarian, and Moore may have recognised this to an extent, but such nuances did not survive Moore's pivotal reformulation of utilitarianism that would shape the progress of the doctrine in the early twentieth century.

Bertrand Russell was also guilty of failing to acknowledge the extent to which he was indebted to Sidgwick. Russell was not considered to be a moral philosopher

consider Sidgwick either to have been part of that camp. Other theorists have also presented Sidgwick and Moore as having a kind of common front in the form of their style of intuitionism, and – more importantly – when it comes to the concept so central to Moore of the good being 'undefinable'. Bernard Bosanquet, who reviewed *Principia Ethica* for *Mind* (Vol.13, 254-261) in 1904, stated at the time that it was clear throughout the book how strongly Moore was influenced by Sidgwick. Skorupski also describes Sidgwick and Moore together as being "leading figures in a reaction against that mood that developed at the turn of the century. The nature of this reaction is not quite easy to catch. Its essence, whether in ethics or in logic or epistemology, is to place great stress on the self-standing autonomy of these subjects as intellectual disciplines...whose basic principles can only be known by 'a priori Intuition'" (Skorupski, 'Spencer and the Moral Philosophers', in, Mark Francis and Michael W. Taylor (eds.) *Herbert Spencer; Legacies* (Abingdon-on-Thames: Routledge: 2011) p. 144. See also Thomas Hurka, 'Moore in the Middle', in, *Ethics*, Vol.113 (2003) 599-628. p. 600). Margolis is particularly specific that what Moore had said on this subject was already present in Sidgwick ('G.E. Moore and Intuitionism', p.37). Decades later, Bernard Williams would still draw the same comparisons, on the basis of Sidgwick and Moore's shared intuitionism (Bernard Williams, 'The Point of View of the Universe', in, Bernard Williams, *Making Sense of Humanity and Other Philosophical Papers 1982-1993* (Cambridge: Cambridge University Press: 1995) p.155). Even E.E. Constance Jones, with her staunch devotion to Sidgwick, emphasised the similarities between the two philosophers, despite Moore's hostility towards Sidgwick. See Jones, 'Mr Moore on Hedonism', in *International Journal of Ethics*, Vol.16, No.4 (1906) 429-464. Thomas Hurka agrees that *Principia Ethica* was certainly highly influential on the development of modern moral philosophy, but disagrees that many of its claims were new, or original ('Moore in the Middle', pp. 599-601). The most important point on which Hurka disputes Moore's originality is his claim that Sidgwick was the first philosopher to state that the good was undefinable. Despite Moore's agreement with Sidgwick on the nature of good, Moore did finally depart from his former teacher's arguments on that eternally problematic issue of hedonism. On Moore's view, 'good' must be a universal property, and thus cannot be defined as 'pleasure', which is a naturalistic description. Moore mainly considered Sidgwick's account of the good to avoid such terms, but according to Moore's strict criteria for how *not* to define the good, Sidgwick was still guilty of equating 'good' with some kind of consciousness.

¹⁶⁷ Hurka, 'Moore in the Middle' pp.602-603

in the traditional sense, as Moore was¹⁶⁸. He was, however, a very popular writer and his highly logical arguments were well known in non-philosophical and public spheres. At least some of these qualities, both personal and intellectual, Russell owed to having been a pupil of Sidgwick's. Still an adherent of the Idealist tradition himself, Russell renewed the by then declining interest in *ME* by attempting to solve the DoPR, and reaching the non-cognitivist conclusion that it is difficult to establish a true reason for acting morally¹⁶⁹. Although this went no way towards aiding the DoPR, it was Russell's preoccupation with this subject that led him to finally reject Absolute Idealism in favour of what was to become his famous idea that knowledge was only to be had at face value. Russell's epistemological emphasis on logic as the only medium through which "direct truths" could truly be known quickly became popular as the "new philosophy of logic", which included an emphasis on the use of scientific approaches to ethics¹⁷⁰.

Russell admitted some years later that both he and Moore were far more indebted to Sidgwick than either of them had admitted, and that they had not given him the respect – either personally or academically - he deserved¹⁷¹. Sidgwick's

¹⁶⁸ Russell's thought did align closely with Moore's on the issue of the non-naturalistic properties of good, which of course was an idea Moore shared with Sidgwick. But it is also arguable that to some extent, it was Moore who was influenced by Russell's own account of the errors of naturalism. Russell published his ideas on the matter in "*The Relation of What Ought to be to What Is, Has Been, or Will Be*" (which was actually a paper he wrote for Sidgwick) in 1893, a full ten years before Moore's *Principia Ethica* appeared. This is not surprising however, since both were pupils of Sidgwick, and Sidgwick had already outlined this argument.

¹⁶⁹ Russell originally theorised that the DoPR might be solved via McTaggart's argument that self-interest and morality could be united by looking towards a day when the Idealist 'Absolute' would become an explicit reality (Russell, 'On the Foundations of Ethics', in, Charles Pigden (ed.), *Russell on Ethics* (London: Routledge: 1999), pp. 206-11). Eventually however, Russell became persuaded of Bradley's non-reality version of the Absolute, and finally stated that he did not believe Practical Reason and self-interest could ever be truly reconciled.

¹⁷⁰ See Russell's 'Knowledge by Acquaintance and Knowledge by Description', in *Proceedings of the Aristotelian Society*, Vol.11 (1911) 108-128; Russell, *The Problems of Philosophy* (Oxford: Oxford University Press: 1912); Russell, *Human Knowledge: Its Scope and Limits* (London: George Allen and Unwin: 1948), among others. (For the secondary literature, see Paul Hager, *Continuity and Change in the Development of Russell's Philosophy* (London: Kluwer Academic Publishers: 1994); A.D. Irvine, 'Russell on Method', in, Godehard Link (ed.) *One Hundred Years of Russell's Paradox* (Berlin and New York: Walter de Gruyter: 2004), and, Irvine, 'Epistemic Logicism and Russell's Regressive Method', in, *Philosophical Studies*, Vol.55 (1989) 303-327). Russell, more than Moore, was the leading figure in this new foray away from classical philosophy and into more obviously analytical philosophy. His extensive list of publications reflects the early decades of the 1900s' fashionable preoccupation with science, and realism. There are, however, grounds on which it is more than reasonable to argue that Sidgwick himself was trying to reintroduce an element of the scientific method into his own approach to sorting out the many ethical theories of his time. He uses the specific language in several places throughout *ME* (see *ME*:vi, I:I:1,-2; III:I:215; III:XI:338, 355, 360; III:XIII:374).

¹⁷¹ "We called him 'old Sidg'", Russell recalls in *My Philosophical Development* (London: George Allen and Unwin: 1959), "and regarded him as merely out of date". As Roma Hutchinson suggests, the

investigations and means of inquiry were also comprehensively logical, and there is little doubt that Sidgwick's work set Russell's early agenda¹⁷². But it is also clear that ME had become something of a relic in this time – a dry and dated work that only served to summarise, however thoroughly, those modes of moral philosophy that were rapidly going out of fashion. Unfortunate though it was, Moore and Russell's attitudes were directly representative of the lasting impression that was attached to Sidgwick as his era faded away.

Russell also maintained the impression that Sidgwick was 'the last surviving representative of the Utilitarians'. When H.A. Pritchard published his well known paper 'Does Moral Philosophy Rest on a Fundamental Mistake'¹⁷³, which questioned those attempts to give answers to the 'why?' of moral obligation and untangled the Moorean association between 'ought' and 'good', he brought back to the forefront of moral philosophy the idea of the 'right'. Utilitarianism thus faced a serious challenge, and Sidgwick was still being held up as its last great defender. In this cacophonous environment of analytical philosophy and the subsequent dwindling faith in utilitarianism, Sidgwick's voice was almost entirely lost.

1.2. Sidgwick in the Mid-Twentieth Century

During the First World War, and the decade that followed it, moral philosophy was relatively quiet. A sort of societal utilitarian attitude had been encouraged in the nation during the war, as the means through which the war effort – and the tragedy - could be supported, and accepted. In the decades that followed however, there were several major resurgences in both the deontological and utilitarian schools. Although

general attitude of these more progressive of Sidgwick's pupils could be seen as one of 'mean-mindedness and contempt' (Hutchinson, Index to *The Collected Stories of Bertrand Russell*. Russell: The Journal of Bertrand Russell Studies, [S.I.], v. 22, jun. 2002. ISSN 1913-8032. Available at: <<https://escarpmentpress.org/russelljournal/article/view/2017>>. Date accessed: 27 Jun. 2016. See also Ronald Beanblossom 'Russell's Indebtedness to Reid', in, *The Monist*, Vol.61, No.2 (1978) 192-204, and P.A. Schlipp (ed.) *The Philosophy of Bertrand Russell* (The Library of Living Philosophers Vol.5) (Evanston and Chicago: Northwestern University: 1944). At a more professional level, Russell and Sidgwick did actually maintain a decent working relationship, as is clear from their correspondence and from Sidgwick's comments on the numerous papers Russell wrote for him. See Kenneth Blackwell, Andrew Brink, Nicholas Griffin (eds.) *The Collected Papers of Bertrand Russell Vol.1* (London: Routledge: 1983) pp.209-10. They also knew each other personally through the Cambridge Apostles, and the Ethical Club.

¹⁷² Russell also departed from Moore's analytic version of consequentialism on the basis of a rather more Sidgwickian caution regarding how to derive the moral 'ought'.

¹⁷³ In, *Mind* Vol.21 (1912) 21-37

in many ways this rising contention furthered the divide between deontological and utilitarian ethical systems, there remained a point of reference that was common to both sides of the debate. This was the work of Henry Sidgwick. In these sections, some of the writers whose theories were held up as rivals to Rational Benevolence in the thesis introduction appear again, but this time in their capacity as figures who contribute something important to the unfolding story of interpretations of Sidgwick's work.

1.2.a. Sidgwick and W.D. Ross

The 1930s saw the beginning of a significant revival in deontological ethics – in the capable hands of W.D. Ross. As was seen in the introduction, Ross offers a moderate form of deontology that incorporates both right and good in such a way that both deontological and utilitarian concerns are to some extent accounted for. For the intuitionism that informs this theory, Ross directly followed Sidgwick.

It is interesting that Ross, as a deontologist, would draw on Sidgwick's intuitionism, when for Sidgwick it had led to (apparently) utilitarian conclusions. But Ross's intuitionism took the form of the non-empirical, *a priori* knowledge type of intuitionism, a tradition that had come down through Thomas Reid and William Whelwell, and had been so convincingly formulated and redressed by Sidgwick¹⁷⁴. Ross seems to have simply ignored the fact that Sidgwick's epistemological application of intuitionism yielded utilitarian results¹⁷⁵. In fact, where Sidgwick

¹⁷⁴ This was as opposed to empiricist intuitionism which argues that moral truth can be known through some kind of moral 'sense'. Many writers unite Sidgwick and Ross on the basis of their shared intuitionism (see for example Hurka, 'Sidgwick on Consequentialism and Deontology', p.133, and Elizabeth Tropman, 'Renewing Moral Intuitionism', in, *Journal of Moral Philosophy*, Vol.6, No.4 (2009) 440-463). John Rawls unites Ross and Sidgwick on the basis of their shared beliefs about the nature and function of conscience (Rawls (Barbara Herman (ed.) *Lectures on the History of Moral Philosophy* (Cambridge, MA: Harvard University Press: 2000) p.77). Interestingly however, John Deigh ('Sidgwick on Ethical Judgement', in Schultz, *Essays on Henry Sidgwick*, pp.241-258) points out that Sidgwick maintained Kant's view that ethical judgement (i.e. in their case the recognition of something rational as right) is a motivator to action, where deontologists such as Ross and Prichard did not.

¹⁷⁵ A possible explanation for the deontological/utilitarian intuitionism cross-over between these two writers is that it was only on the *type* of intuitionism that each upheld that they were divided. Originally, the term 'intuitionism' was synonymous with moral pluralism: It was only during the 1700s and 1800s that the word became associated with the form of epistemological intuitionism that dominated the early 1900s. Sidgwick was considered a monist in that he defended the principle of utilitarianism *only*, where Ross defended a pluralist account of several deontological *a priori* moral principles that are ranked in terms of priority. There is then an apparent division between Ross and Sidgwick over the issue of moral pluralism, but Ross does not seem to have recognised it.

appears in Ross' defence of deontology in *The Right and the Good*, there is considerable weight attached to the points. Most significantly, during Ross' refuting of Moore's understanding of 'rightness' (as productive of the greatest possible good), he draws on Sidgwick's argument that Bentham had never really given a definition of 'the right' that wasn't in the end basically a tautology¹⁷⁶. Ross then utilises Sidgwick's point that "the meaning of 'right' or 'ought' is in fact 'too elementary to admit of any formal definition', and explicitly quotes Sidgwick's express repudiation of 'the view that 'right' means 'productive' of any particular sort of result'¹⁷⁷. This is what informs Ross's intuitionist view that "right' is an irreducible notion'¹⁷⁸. And *this* is what informs the foundational concept of Ross's *deontological* ethics.

Sidgwick's direct relevance to Ross' deontology here is striking – as is Ross' overlooking of the fact that it was a utilitarian whose notion of rightness he agreed with. That particular relevance, of Sidgwick's epistemology to deontological concepts, was once again lost in the shifting tides of new philosophy.

1.2.b. Sidgwick in Act and Rule Utilitarianism

Deontology advanced significantly under Ross. *The Right and the Good* was generally considered to be one of the most outstanding treaties on moral theory of the twentieth century, and utilitarianism suffered a sudden and brutal decline in credibility in the wake of the events of the Second World War. Not long after this however, utilitarianism was to receive one of its most important reformulations; the 'rule'/'act' distinction, defended mainly by Richard Brandt in his 1963 paper, 'Towards a Credible Form of Utilitarianism'¹⁷⁹.

¹⁷⁶ *The Right and the Good*, p.7

¹⁷⁷ *The Right and the Good* p.8. Ross, here, reads Sidgwick as I do in that passage from *ME*. What is right, for Sidgwick, is not dependent on the production of results. i.e. his notion of rightness is non-consequentialist. This reinforces my argument for the presence of a non-consequentialist type of duty element in Sidgwick's work, and it is supported even further within this context of Ross' work, by the fact that Ross' system depends almost entirely on the notion of non-consequentialist duty.

¹⁷⁸ *The Right and the Good* p.12

¹⁷⁹ In, Michael D. Bayles (ed.) *Contemporary Utilitarianism* (New York: Doubleday: 1968), pp.143-186. The title of Brandt's paper is also revealing of the paper's purpose. As opposed to the traditional utilitarian position, which generally holds that the principle of utility is to be applied on case by case basis in order to produce the most utility (or 'happiness', to use Bentham's understanding), rule

There are noticeable differences between Brandt and Sidgwick's versions of utilitarianism, and Brandt did not often reference Sidgwick directly. He did, however, label Sidgwick as a rule utilitarian¹⁸⁰. This led to a highly interesting debate over the nature of Sidgwick's work, headed up by fierce rule-utilitarian critic, J.J.C. Smart. Smart firmly declared Sidgwick to in fact be an act-utilitarian, in his famous denunciation of the existence of any such thing as 'rule utilitarianism'¹⁸¹. Both views

utilitarianism argues that the most beneficial results are brought about by following certain impartial *rules* that are thought to maximise utility, rather than by individual actions. The rightness or wrongness of an action is therefore judged not according to the results that it brings about immediately, but by its conformity to a rule that has been judged to bring about the best results (see also Richard T. Garner, and Bernard Rosen, *Moral Philosophy: A Systematic Introduction to Normative Ethics and Meta-ethics* (New York: Macmillan:1967), p.70). Those rules are what determine moral rightness. In this way, rule utilitarianism aimed to avoid the common accusation advanced against utilitarianism in its classical form, that its lack of a proper objective concept of 'right' was ethically unsound, and socially and morally dangerous.

See also Brandt's *Ethical Theory: The Problems of Normative and Critical Ethics* (Englewood Cliffs, N.J: Prentice-Hall: 1959), *A Theory of the Good and the Right* (Oxford: Clarendon Press: 1979), and *Morality, Utilitarianism and Rights* (Cambridge: Cambridge University Press: 1992). In the 1963 paper, Brandt looks to J.O. Urmson and his interpretation of Mill as a rule utilitarian. But some writers, such as Michael McDonald, have drawn parallels between Sidgwick's work and Brandt's, especially where *A Theory of the Good and the Right* is concerned (see McDonald, 'Critical Notice', in, *Canadian Journal of Philosophy*, Vol.12, No.2 (1982) 375-389). McDonald points out that Brandt seems to share Sidgwick's conviction that morality is inherently connected to rationality, and that both philosophers advocate a form of 'desire theory', wherein we act according to what we feel will make our lives go best) (pp. 381-382). It is indeed the case that Brandt and Sidgwick both anchor their ethical theories in human rationality, but McDonald also points out that Brandt does ultimately depart from Sidgwick by rejecting Sidgwick's Philosophical Intuitionism. Brandt is careful to avoid the 'normative generalisation' of the argument – defended by Sidgwick – that there is a basic premise (derived by intuitionism or some other abstract means) that it is best to act so as to maximise desires and that therefore actions that conform to this are the best actions. This, Brandt says, is precisely what the reasoning of *A Theory of the Good and the Right* is not meant to suggest (Brandt, *Theory of the Good and the Right*, p.154). Brandt's alternative approach is an empiricist one, that argues – in reverse to the previous argument – that if we act rationally we are likely to satisfy desires, and therefore acting rationally can be recommended to agents, because it conforms with what we already know and do. In this way, Brandt is closer to Mill than he is to Sidgwick, but Sidgwick and Brandt still share the welfare-maximising basis that provided the basis for Brandt's rule utilitarianism.

¹⁸⁰ It was at this point in history that G.E.M. Anscombe introduced the phrase 'consequentialism' in her important paper 'Modern Moral Philosophy' (in, *Philosophy*, Vol.33, No.124 (1958) 1-19. Actually via a specific reference to the work of Sidgwick, Anscombe intended the word 'consequentialist' to refer to the denial that there is any significant moral difference between results of an action that are intended, and those that are foreseen, but *not* intended (p. 10). These days, 'consequentialism' is often used as a synonym for utilitarianism, but this is inaccurate: Anscombe intended the word to distinguish between the attitude of classical utilitarianism that appears to hold individuals to account for unforeseen bad consequences of actions, and that of utilitarianism as it was reformulated by Sidgwick – and thus continued after him - in which one cannot be held responsible for the bad consequences of one's actions, even if they were foreseen, provided that they were not the agent's intention.

¹⁸¹ See Smart, 'Extreme and Restricted Utilitarianism', in, *The Philosophical Quarterly*, Vol.6, No.25 (1956) 344-354; Smart, *An Outline of a System of Utilitarian Ethics* (Carlton: Melbourne University Press:1973); Smart and Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press: 1973); David Lyons, *Forms and Limits of Utilitarianism* (Oxford: Clarendon Press: 1965). Both Smart and Lyons denounced 'rule' utilitarianism on the basis that it is simply not a viable form of the doctrine. In his well known paper of 1956 (in which he refers to 'extreme' and 'restricted' utilitarianism,

are understandable. On the one hand, Sidgwick could be understood to be arguing that specific ‘rules’ are to be derived from more generalised principles – which seems to be the route he takes later in ME¹⁸². On the other hand, Sidgwick defends Benevolence as the true end of moral duty, by which he means the obligation to regard the good of any other individual as much as his own, and this appears distinctly *act* utilitarian if that form is understood as focusing on the effects of individual actions, as opposed to the *type* of actions done, and how far they conform to a rule. It is this to which Smart confidently directs the reader when he argues that Sidgwick is an act-utilitarian and therefore one who avoids the certain collapse into rule-worship, and deontology¹⁸³. Today, the general consensus is that he was an act utilitarian – but that his act utilitarianism was a less direct version than Smart’s, for example¹⁸⁴. A great many others, however, have not been so sure that Sidgwick classifies quite this easily¹⁸⁵. The act/rule Sidgwickian debate is in this way reflective

ahead of Brandt’s use of the definitions ‘act’ and ‘rule’), Smart points out that to adhere firmly to such rules is actually just a form of ‘superstitious rule worship’, more akin to a deontological position – and yet if breaking of the rules is permitted, those rules are no longer relevant anyway. Restricted utilitarianism thus ‘collapses’ into extreme (or ‘act’) utilitarianism.

¹⁸² “[philosophical intuitionism] ...while accepting the morality of common sense as in the main sound, still attempts to find for it a philosophic basis which it does not itself offer: to get one or more principles more absolutely and undeniably true and evident, from which the current rules might be deduced....” (I:VIII:103). During his discussion of the method of utilitarianism, Sidgwick clearly writes from the point of view that there are rules from which utilitarian-type actions are derived (IV:V:488-91).

¹⁸³ See also William H. Shaw, who also asserts conclusively that Sidgwick was an act-utilitarian (Shaw, *Contemporary Ethics: Taking Account of Utilitarianism* (Oxford: Blackwell: 1999) p.10

¹⁸⁴ See John Gray, ‘Indirect Utility and Fundamental Rights’, in, Gray, *Liberalisms: Essays in Political Philosophy* (London: Routledge: 1989); Bernard Williams, *Making Sense of Humanity and other Philosophical Essays 1982-1993* (Cambridge: Cambridge University Press: 1995); Hooker, *Ideal Code*, p.143; Singer and Lazari-Radek, *Point of View of the Universe*, p.176; Roger Crisp, *The Cosmos of Duty: Henry Sidgwick’s Methods of Ethics*, (Oxford: Oxford University Press: 2015) pp.206-214; Robert E. Goodin, *Utilitarianism as a Public Philosophy* (Cambridge University Press:1995) p.16. Geoffrey Scarre has recently utilised Sidgwick to argue *against* act-utilitarianism, pointing out that the selfless optimisation that seems to be required by act utilitarianism (and the sort Scarre attributes to Peter Singer) leads to ‘paradoxically narrow-minded’ remote and unfeeling individuals, and that Sidgwick did not believe that one would be motivated to promote another’s utility, if he was required to relinquish his own (Scarre, ‘Is Act Utilitarianism the ‘Ethics of Fantasy’, in, *Society for Applied Philosophy* Vol.15, No.3 (1998) 259-270, p.262.

¹⁸⁵ See R.G. Frey, ‘Act Utilitarianism: Sidgwick, Bentham, or Smart?’, in, *Mind*, Vol.86, No.341 (1977) 95-100; Hooker, ‘Sidgwick and Common Sense Morality’, 355-360; Rawls, ‘Two Concepts of Rules’, in, *The Philosophical Review* Vol.64 (1955) 3-32, p.14; John C. Harsanyi, ‘Rule Utilitarianism and Decision Theory’, in, *Erkenntnis*, Vol.11, No.1 (1977) 25-53, and ‘Morality and the Theory of Rational Behaviour’, in, Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond*, (Cambridge: Cambridge University Press) 56-61.

Frey’s article is particularly illuminating of this debate. From the first page it is clear that Frey is convinced that Sidgwick is an act-utilitarian: His phrase “...it can often be used to provide support for parts of what he calls the morality of common sense...” (p.95) is a perfect example of the function of act-utilitarianism, as opposed to the reverse situation in which rules might be used to restrain simple utilitarian actions. However, in the following sentence, Frey specifically refers to the presence of this

of Hayward's early words – that Sidgwick as a moral philosopher is often very hard to classify at all¹⁸⁶.

1.2.c. Sidgwick and William Frankena

In 1963, William Frankena produced his seminal deontological work, *Ethics*¹⁸⁷. Like Ross, Frankena's semi-reconciling (of deontological and utilitarian principles) theory was discussed in the thesis introduction for its relevance to the theory of Rational Benevolence. But again like Ross, Frankena is highly important in this narrative for where his deontology also made significant reference to Sidgwick.

Though little secondary literature draws attention to the fact, Frankena shares many characteristics with Sidgwick, both academic and personal. Frankena's aim in *Ethics* was much the same as Sidgwick's in *ME* in that Frankena sets out to examine each ethical position, such as egoism, hedonism, and relativism, in turn (although Frankena does state a more specific intention to go beyond the meta-ethics to

reverse situation in Sidgwick's writing as well: "...act-utilitarianism is [not] going to be allowed to sweep aside just any of the views and convictions of the 'plain man'..." Frey concludes here that Sidgwick is arguing that "act utilitarianism will *sometimes* give results that warrant amending common sense morality, and *at other times* will give results that warrant amending act-utilitarianism....the point here is not *which* amendments Sidgwick was in favour of, but *that* he thought that amending the theory was an appropriate response in the light of at least some deeply counterintuitive results". That is, Frey obviously considers Sidgwick to at least be a more moderate or conservative act utilitarian than Smart. Even more revealing is the argument that the second situation that Frey describes could be interpreted as representing Sidgwick to essentially hold a fully rule utilitarian position. Frey does not retract his definition of Sidgwick as an act-utilitarian, but Frey's particular portrayal of Sidgwick's act-utilitarianism lends further support to the idea that Sidgwick could be either an act or a rule utilitarian, certainly in comparison to either Smart or Bentham.

Incidentally, I agree that this debate over whether Sidgwick was an act or a rule utilitarian should exist, given that it is not truly clear in the text of *ME*. However, as it is one of my specific aims in this thesis – and certainly in this chapter – to dilute the image of Sidgwick as a utilitarian, I have not specifically engaged with this aspect of Sidgwick studies.

¹⁸⁶ At about this time, A.R. Lacey, in 'Sidgwick's Ethical Maxims' (*Philosophy*, Vol.34, No. 130 (1959), 217-228) argued that Sidgwick did not distinguish between 'strictly non-teleological duties, such as the Kantian duty of absolute veracity, and semi-teleological ones' (p. 218). Furthermore, Lacey believed there to be a strong sense of 'duty' in Sidgwick, as shown by the fact that Lacey was able to identify 4 'duty premises' in Sidgwick's work (p. 220). (Only the first of these – the duty to produce good – is positive. The other three are negative statements about what is *not* our duty). Ultimately, Lacey believes Sidgwick's notion of duty is a teleological one. This is a defensible position, as there is no doubt that there is a teleological element in Sidgwick's work. I dispute it however, on the basis that I believe Sidgwick's interpretation of rightness (as being that which is ultimately reasonable) to call into question whether Sidgwick's theory of moral obligation is *entirely* teleological (See chapter 3). Lacey is important here for sustaining an undercurrent of recognition that duty of some kind was a major focus of *ME*.

¹⁸⁷ Frankena, *Ethics* (Englewood Cliffs, N.J: Prentice-Hall: 1963)

“providing the general outlines of a normative theory”, and his approach to the various ethical positions was always one of thought and consideration, never of dismissal¹⁸⁸. Interestingly, it could also be argued that Frankena too suffers a similar fate to Sidgwick’s in having been typecast as a certain kind of moralist when actually his theory far surpasses the boundaries of that label. For, where Sidgwick is almost universally considered to be a utilitarian despite his clear sympathy for deontological arguments, Frankena is almost universally considered to be a deontologist, despite his clear respect for utilitarian aspects of morality.

Most importantly for the relationship between the two, Frankena frequently references Sidgwick – and often within the *specific* context of deontology. There are numerous examples of this. The first source to which Frankena recurses when giving an example of the abstract rules that might be held by a ‘pure rule-deontologist’, is Sidgwick, and Sidgwick’s Principle of Justice – and he actually does this twice¹⁸⁹. Frankena goes on to directly compare Kant’s isolating of a single basic deontological principle with Sidgwick’s own account of the abstract principle of Justice (as Sidgwick intended) and later again groups Sidgwick together with Kant in his discussion of what is good as an end¹⁹⁰. Here, Frankena specifically states that Sidgwick combines a hedonistic theory of value with a deontologist (and not a teleological) theory of the duty of obligation. Similarly to Ross, it is Sidgwick’s notion of right that makes this possible: ‘One may adopt a hedonistic theory of value without adopting any such theory of obligation’, Frankena says, ‘A hedonist about the good may be a deontologist about the right; roughly speaking, Butler, Kant, and Sidgwick combine hedonism about the former with deontologism about the latter’¹⁹¹.

The importance of these words from Frankena cannot be overstated. For the first time in the history of Sidgwick’s appearances in the work of other philosophers, Frankena explicitly recognises not just the presence of the deontological aspect of Sidgwick’s theory, but also the fact that this deontological aspect contributes as significantly to his moral theory as the utilitarian aspect. If anything, the fact that

¹⁸⁸ Ibid, p.5

¹⁸⁹ Ibid, pp.15, 25

¹⁹⁰ Frankena states later that Sidgwick’s principle of justice is not actually sufficient as a normative ethical principle, given that it only states that we should act according to rules if we want to be just, but tells us nothing about what the rules are to be (*Ethics*, p.39). But Frankena is still discussing Sidgwick within the context of that single principle which Frankena earlier identified to belong to the realm of the pure rule deontologist.

¹⁹¹ *Ethics*, p.68

Frankena states Sidgwick to be a deontologist about the right might seem to give more authority to deontology in Sidgwick's system than utilitarianism¹⁹². As if to add further weight to his point, Frankena includes *Kant* in the same category. There could hardly be a more direct example of the fluidity of Sidgwick's work for either occupying a space between deontological and utilitarian theories, or for being able to move between them, than how he appears in this surprisingly brief passage in Frankena.

Although they are separated by some thirty years, it seems appropriate to group Ross and Frankena together here. Both Ross and Frankena are deontologists. Both had noticed that Sidgwick's concept of the right was not a utilitarian one. Both follow Sidgwick in this concept of rightness. The view of neither Ross nor Frankena has prevailed. It is commonly recognised that both philosophers reference Sidgwick, but the full significance of *how* and *why* they reference him remains, it seems to me, unfulfilled¹⁹³.

1.2.d. Sidgwick and John Rawls: Sidgwick's Appearance in Political Deontology

There has already been, throughout this history of Sidgwick's place in moral philosophy, an enormously wide variety of interpretations of his work and thought.

¹⁹² Frankena is in clear agreement with my argument in this thesis that Sidgwick's interpretation of the right is a deontological one, understood as representing non-consequentialist duty. But Frankena is also the only other scholar to specifically reference the fact that Sidgwick's theory of obligation may not actually be teleological. I agree entirely with Frankena that Sidgwick's argument for the right poses a challenge to the teleological aspect of his work (see chapter 3).

¹⁹³ Alasdair MacIntyre's comments on Sidgwick are typical of how Sidgwick was viewed for so long during the mid twentieth century. In *Short History*, MacIntyre identifies Sidgwick as a melancholy sort of figure, whose 'defects are usually the defects of his age' (p. 235). By this MacIntyre means that Sidgwick was preoccupied with his own crises of philosophy and faith, and stood at the brink of the clash between intuitionism and the utilitarian consciousness that had dominated his age. Whereas both of these claims are true, MacIntyre falls into the same trap as many twentieth century theorists by simply placing Sidgwick at the end of the classical era, and then closing the door upon him. (Marcus Singer is explicit that this is a misunderstanding of Sidgwick on the part of MacIntyre. See Singer (ed.) *Essays on Ethics and Method* p.xxxiv-xxxv). MacIntyre also fails to draw on Sidgwick in his own, influential argument for virtue ethics that rationality is fundamentally flawed, despite the clear similarities between this claim and Sidgwick's. In his 'Whose rationality?' theory, MacIntyre argues that human rationality is not constant: It is empirical that what is considered 'rational' varies from culture to culture, from individual to individual. Rationality cannot then be held up as the paradigm of ethics (MacIntyre, *After Virtue* (Notre Dame, IN: University of Notre Dame Press: 1981), pp.6-22; *Whose Justice? Which Rationality?* (London: Duckworth: 1987) pp.1-11. Whereas the parallel between MacIntyre's claim and Sidgwick's own primary concern that reason was full of conflicts may seem to be apparent, MacIntyre's only reference to Sidgwick is that he finally proved the teleological framework of ethics to have failed (*After Virtue*, pp.64 – 65). Sidgwick is then swiftly dispensed with.

But perhaps the most surprising use of Sidgwick is that found in the work of American philosopher John Rawls during the political upheaval of the 1970s.

The 1970s was an interesting era for moral philosophy¹⁹⁴. Ethical debate was changing in accordance with the rapidly changing socio-political context, which was characterised by a rise in socially progressive values such a general political awareness and the subsequent rise in interest in the concept of civil and political rights¹⁹⁵. Together, these themes were instrumental in the increasing support for egalitarian and individualistic social values. Where utilitarianism was unlikely to be able to meet such demands, the natural association between the concept of civil rights and the absolute 'right' inherent in deontological approaches to ethics was obvious. New trends in moral debate were about to rise to meet the new ethical demands - and this was the sphere in which Rawls was working.

Rawls' revitalising of Kant's doctrine as against utilitarianism in his seminal *Theory of Justice* (1977) was nothing short of revolutionary. His deontological theory of 'justice as fairness' presents a firm case for the prioritising of the right over the

¹⁹⁴ It is arguable that as the decades had pressed on, both deontology and utilitarianism in their classical forms had come to a close. The 'divide' between the two positions, however, raged on. In 1967, Philippa Foot gave a now famously popular argument against consequentialist theories of morality in the form of her 'Trolley Problem', a thought experiment that involves the reader deciding whether to allow five workmen on a track to die under the wheels of a runaway trolley by simply allowing the trolley to run on, or whether to intentionally turn to the wheel onto another track, which will kill only one man (Foot, 'The Problem of Abortion and the Doctrine of Double Effect', in, *Virtues and Vices and other Essays in Moral Philosophy Vol.19* (Oxford: Blackwell Publishers: 1978). The utilitarian view, Foot argues, not only makes the second choice morally permissible, but actually morally *obligatory* – that is, utilitarianism makes it morally obligatory to intentionally kill a person, in certain circumstances. The Trolley Problem was also extensively analysed – and supported - by Judith Jarvis Thompson ('The Trolley Problem', in, *The Yale Law Journal*, Vol. 94, No.6 (1984) 1395-1415). Warren Quinn also followed Foot in his argument that the difference between killing or letting die came down to a difference between positive and negative rights (Quinn, 'Action, Intentions, and Consequences; The Doctrine of Doing and Allowing', in, *Philosophical Review* Vol.98, No.3 (1989) 355-382). Despite this, R.M Hare's preference utilitarianism was leading the pro-utilitarian side, and the pure deontology of Kant and Ross seemed to have been diluted by Frankena's pointing to the importance of utilitarian principles. G.E. Anscombe's influence was also being felt by this time; a combination of her attack on both deontology and utilitarianism, and Frankena's own leaning towards the field of virtue ethics in his discussions of moral character, finally allowed virtue ethics onto the field as a real viable alternative to the two schools of thought that had so dominated ethical debate until this point. In terms of deontology/utilitarianism relations however, this stage in the history is an example of old habits dying hard. However many modifications had been made to either doctrine, wherever those two labels could be identified they were still considered to represent two fundamentally conflicting ideas.

¹⁹⁵ This was particularly prevalent in America. Anti-nuclear weapons protests, the rise of feminism and the change in the role of women, and widespread withdrawing of public support for the Vietnam war are all examples of the atmosphere that was developing (sometimes with ironic violence) during this time – the atmosphere of freedom of expression, and of personhood. Novelist Tom Wolfe described the combination of these factors as having made the 1970s into 'the 'Me' decade', in his influential essay of 1976, 'The 'Me' Decade and the Third Great Awakening', in *New York Magazine*, (New York: Condé Nast) 23rd August.

good, and advances a sophisticated case against utilitarianism¹⁹⁶. In the notes I have given a highly redacted version of what is an extremely complex and multi-layered theory, but for present purposes, it is not inaccurate to say that in Rawls we have – on Rawls’ own definition – a classically deontological system that gives absolute precedence to the right¹⁹⁷. But even in this most important of neo-Kantian/anti-utilitarian theories, its author saw fit to recourse frequently - and often positively - to Sidgwick¹⁹⁸.

Sidgwick appears in Rawls’ work in multiple ways. Firstly, Rawls recognised the importance of the utilitarian doctrine *only* as it had been expounded by Sidgwick - and accepted it as the only really serious alternative to his own theory of justice as fairness¹⁹⁹; Rawls’ engagement with utilitarianism throughout his seminal *A Theory*

¹⁹⁶ Rawls identifies the utilitarian position as being the principle that an individual will advance as far as possible his own welfare/good, and that when applied to the group of society, the aim is to advance as far as possible the welfare/good of the group (Rawls, *A Theory of Justice*, pp.19-21, 294-294-295, 394-396 ; Rawls, *Political Liberalism* (New York: Colombia University Press: 1993 p. 21). From the perspective of Rawls’ political philosophy, the problem with this is most clearly demonstrated when individual case utilitarianism and social utilitarianism are viewed congruently. Because the aim in both cases is simply the satisfaction of desires, how this satisfaction is brought about does not matter; that is, without an independent account of the right “there is no reason in principle why the greater gains of some should not compensate for the lesser losses of others; or, more importantly, why the violation of the liberty of a few might not be made right by the greater good shared by many” (*Political Theory* p.23). Rawls’ main claim against utilitarianism therefore is summarised as ‘it does not take seriously the distinction between persons’ (*Theory of Justice*: 73-93; *Political Liberalism*, p.24). The good simply cannot be given the priority over the right (See Will Kymlicka, ‘Liberalism and Communitarianism’, in, *Canadian Journal of Philosophy*, Vol.18, No.2 (1988)181-203, p. 184). Instead, Rawls gives an account of the right that is independent of the good. In order to argue for this, Rawls constructs a type of thought experiment that is based on the concept of a social contract, whereby all parties within the contract agree on an arrangement to which everyone can mutually agree. This is the ‘original position’, a hypothetical situation in which everyone will be a member of a society that is to be created, but does not know anything about what their role in that society might be (*Theory of Justice*, pp.15-19, 104, 145, 587). They will be entirely ignorant as whether they are wealthy or poor, male or female, a minority or a majority etc. This being the case, and humans being rational, Rawls argues that the first logical desires of these people will be fairness, freedom and equality – or, as Rawls terms it, the principle of justice. Following Kant’s lead, Rawls argues that it would simply be irrational to propose anything else. In Rawls’ words, “my aim is to present a conception of justice which generalises and carries to a higher level of abstraction the familiar theory of the social contract...the guiding idea is that the principles of justice for the basic structure of society are the object of the original agreement. They are principles that free and rational persons concerned to further their own interests would accept...as defining the fundamental terms of their association (*Theory of Justice* 1971, p.10)

¹⁹⁷ *Political Liberalism*, p. 26.

¹⁹⁸ Rawls had enormous respect for *The Methods*, which he considered to be a better representation of the utilitarian doctrine than Mill’s (*A Theory of Justice*, p. 20). For Rawls, Sidgwick represented both a model example of how to go about investigation moral theory in general, and the epitome of the classical utilitarian doctrine, in its “clearest and most accessible formulation” (*Theory of Justice*, p.20).

¹⁹⁹ *Theory of Justice*, p.417

of *Justice* is, essentially, engagement with Sidgwick²⁰⁰. Furthermore, Sidgwick was the only utilitarian whose ethics were relevant for Rawls' strictly deontological system. He referenced Sidgwick's account of the good, and perceived that Sidgwick also advocated a form of deliberative rationality, in which a person's good is determined by what he would choose if all paths and consequences were available to him at the time²⁰¹. Most specifically, Rawls drew specifically on Sidgwick's egalitarian views on formal justice and social class²⁰².

In the secondary literature, one of the most common points on which Rawls and Sidgwick are united is that of reflective equilibrium - the process through which we arrive at the principles generated by Rawls' original position thought experiment²⁰³.

²⁰⁰ See Gerald F. Gaus and Chandran Kukathas (eds.) *Handbook of Political Theory* (London: SAGE Publishing:2004) pp.413-417. Gaus and Kukathas also argue, with great directness, that 'political theorists must take Sidgwick seriously if they take Rawls seriously', and that this is based on Rawls' own theory (p.413). This is to Sidgwick's great credit and potential as a political theorist, as well as a philosopher.

²⁰¹ *Theory of Justice*, p.366. Rawls defines Sidgwick's account of the 'good' as 'the satisfaction of rational desire' (Richard Kraut, *What is Good and Why: The Ethics of Wellbeing* (Cambridge MA: Harvard University Press: 2009) p.124).

²⁰² See *Theory of Justice*, pp.51, 63. Rawls points to Sidgwick's superior avoidance of the 'priority problem' brought about by reliance on intuitionism as the sole method of moral epistemology (whereby considered judgements can possibly be identified, but not weighted appropriately), that Sidgwick achieves through his systematic and highly defensible argument that the principle of utility is that to which we will all ultimately appeal, no matter what our theoretical moral standpoint (*Theory of Justice*, p. 36. This apparently empirical argument is another reason why Rawls afforded such credit to Sidgwick's utilitarianism). Rawls also finds Sidgwick to be in agreement with his argument that "the expectations of those with the same abilities and aspirations should not be affected by their social class" (*Theory of Justice*, p. 63) – a crucial point for the heavy emphasis on the priority of individuals that dominates so much of Rawls' political and moral theory. Rawls agrees with Sidgwick that Kant is unclear why the lives of the saint and the scoundrel are not equally the result of free and rational choices (when Kant has argued that we only become our true selves when we act in accordance with the moral law), and aligns himself with Sidgwick's objection by suggesting the original position as the medium through which we can see which principles free and rational people would choose (*Theory of Justice*, pp. 224-225).

²⁰³ See *Theory of Justice*, pp.18, 15. Several theorists have argued for the similarity between Sidgwick's work and Rawls' reflective equilibrium (see for example Phillip Pettit and Chandran Kukathas, *Rawls: 'A Theory of Justice' and its Critics* (Cambridge: Polity Press: 1980), pp.7-8; B. Barry 'John Rawls and the Search for Stability', in, *Ethics*, Vol.105 (1995) 874-915; Brad Hooker *Ideal Code, Real World* p.9; Frank Snare 'John Rawls and the Methods of Ethics', in, *Philosophy and Phenomenological Research* Vol.36, No.1 (1975) 100-112, p.109); Schultz, *Eye of the Universe*, 196-197). Rawls himself draws the parallel in *A Theory of Justice* when he states that the process of reflective equilibrium is the basis of moral theory and has been for centuries (*A Theory of Justice*, pp.44-45), including in Sidgwick's writing. Peter Singer is very clear that reflective equilibrium has existed throughout most of the history of moral philosophy, and that Rawls directly attributes his inspiration for it to Sidgwick (Singer, 'Sidgwick and Reflective Equilibrium', in, *The Monist*, Vol.58, No.3 (1974) 490-517, p.490), but he does not agree that this is the method that Sidgwick uses. Singer argues that Rawls relies on Schneewind's article 'First Principles and Common Sense Morality in Sidgwick's Ethics' (in, *Archiv Für Geschichte der Philosophie* Vol.45, No.2 (1963) 137-156) but that this too was mistaken: Singer's argument is based on the fact that both these philosophers understood Sidgwick to be 'attempting to bring people over to utilitarianism' above all other moral

Rawls himself draws the parallel in *A Theory of Justice* when he states that the process of reflective equilibrium is the basis of moral theory and has been for centuries, including in Sidgwick's writing²⁰⁴. Rawls and Sidgwick are also similar in their *approaches* to moral theory, though their results on the latter may not have been the same²⁰⁵.

It is possible to overstate the connection, or the nature of the connection, between Rawls and Sidgwick: They are, after all, philosophers of two very different schools of thought, and it is unlikely that they would have finally agreed on what, ultimately, is moral. And yet, theirs is perhaps the most important relationship in the history of Sidgwick's work. Unlike where Sidgwick had appeared in rather more moderate forms of deontology, such as the systems of Ross and Frankena, Rawls cannot be described as presenting a moderate theory. This is the single most comprehensive

theories, when this is not what he is doing. Singer points to Sidgwick's own revising in the preface to the second edition of *ME* of the point that he is *not* aiming to present a system of normative ethics (*ME* p.xx, in, 'Sidgwick and Reflective Equilibrium', p.502). For further points on Singer's engagement with Rawls over the question of Sidgwick's reflective equilibrium, see the section on Singer below.

²⁰⁴ *Theory of Justice*, pp.44-45

²⁰⁵ Schneewind makes this point when he unites Rawls and Sidgwick on the basis of their shared conception of the relationship between general and specific moral principles. According to Schneewind, Rawls and Sidgwick agree that general moral principles are only fully justified if they serve to systemize our more specific moral convictions (Schneewind, 'Korsgaard and the Unconditional in Morality', in, *Ethics* (1998), Vol.109, No.1, 36-48, p. 44) which would account for a large part of the reason that Rawls held Sidgwick's approach to ethics in *ME* in such high regard. Frank Snare also focuses on the similarities between Rawls and Sidgwick's 'methods'. In *John Rawls and the Methods of Ethics*, Snare evaluates the way in which Rawls determines the basic principles of morality. Overall, Snare is arguing that Rawls followed Bentham, Mill and Sidgwick in his own inquiry into normative ethics, citing Sidgwick's phrase "the point of view of the universe", as evidence that the objective viewpoint 'method of ethics' is clearly a concept shared by Rawls' original position theory (*Rawls and the Methods of Ethics*, p.106). Snare appears to have an unusual idea of how Sidgwick was using the word 'method'. Apparently ignoring Sidgwick's appeal to the reader in the preface and the introduction to bear in mind that he is *not* aiming to produce a singular 'method' through which ethics should be done (i.e. his express point is that he is *not* aiming at a system of normative ethics), Snare doesn't agree that utilitarianism and ethical egoism are 'methods' at all, but rather principles that would emerge from some other actual method (also see Schultz, *Eye of the Universe*, p.150 for agreement that "caution" is required when interpreting what Sidgwick meant by 'method'). Snare is right to suggest that Sidgwick's own 'method' is philosophical intuitionism (or at least, that his method of Rational Benevolence resulted from philosophical intuitionism), but is wrong to accuse Sidgwick of having used the word 'method' misleadingly: I believe that it is Snare who has misunderstood Sidgwick's intentions in *ME*. Whereas it might be inaccurate to interpret the entirety of *ME* as a treatise on normative ethics, as Snare seems to have done, it is not wrong to argue that Rawls followed Sidgwick's approach to *how* to do normative ethics. For example, Snare argues that in admitting that some principles may have to be modified in the original position, Rawls is mirroring Sidgwick's argument about the necessary modification of 'considered convictions' (*Rawls and the Methods of Ethics*, p.109). What is even more striking about Snare's article is how readily he draws parallels between Rawls, and Bentham, Mill and Sidgwick, rather than between Rawls and Kant – which he does barely at all. Snare even attributes the 'original position' to Bentham and Mill's ideas about 'proofs' and 'first principles', rather than to the hypothetical situation that is so integral to Kant's categorical imperative.

and decisive defence of modern deontology to have been written to date, and Sidgwick appears in it extensively. Although Rawls stands at the farthest end of moral theory to Sidgwick, he utilises him for many of the points so crucial for that position.

Sidgwick and Rawls' somewhat complex relationship is perhaps best summarised by Claudia Card in her 2002 paper 'Responsibility Ethics, Shared Understandings and Moral Communities', when she states both Sidgwick and Rawls to be proponents of the 'theoretical-judicial' model of ethical theorising (which is characterised by intellectualism, rationality and individualism, among other things)²⁰⁶. Card sees this represented by Rawls' original position, and points not just to the fact that Rawls was influenced by Sidgwick in this respect, but the way in which he influenced him:

"Rawls was strongly influenced by Sidgwick. He did not become a follower or adherent of Sidgwick's position, but rather, took seriously the challenges Sidgwick put forward in *The Methods* and tried to meet some of them"²⁰⁷. Card's words encapsulate Sidgwick's legacy to Rawls, while acknowledging the distance between them²⁰⁸. Overall there is no doubt that Rawls' work has been crucial in maintaining Sidgwick's philosophical (and often political) relevance – and the significance for this of the fact that Rawls was such a staunch deontologist cannot be missed

1.2.e. Sidgwick Revitalised: Sidgwick and Derek Parfit

Rawls frequently used Sidgwick as a model for how ethical inquiry should be done, but he did not draw on Sidgwick's actual moral theory itself. Use of Sidgwick's work was to change distinctly in the next decade however, with a theorist who would become one of the most influential philosophers – and one of the most significant Sidgwick scholars – of the 20th Century, Derek Parfit.

Parfit has an indispensable role in Sidgwick's story, both for revitalising Sidgwick's reputation as a philosopher in his own right, and for making him truly relevant to

²⁰⁶ In, *Hypatia*, Vol.17, No.1 (2002) 41-155

²⁰⁷ Ibid, p.142.

²⁰⁸ Michele M. Moody-Adams is another example of the many scholars who believe Rawls to be one of the writers most integral to the recent revival in interest in Sidgwick's work. (See Moody-Adams, 'Review of Schultz (ed.) *Essays on Henry Sidgwick*, in, *Victorian Studies*, Vol.37, No.1 (1993) 149-150, p.149)

contemporary ethics²⁰⁹. Much of this is the result of the extremely high regard in which Parfit holds Sidgwick's skill as a writer. Where the century before Parfit has often considered Sidgwick to be dry and inaccessible, Parfit takes a very different view, and throughout the preface of his great 2011 work *On What Matters*, he is eager to encourage wider reading of *ME*. He points out that Sidgwick never repeats himself, that he makes important points concisely, and that he writes with clarity, however complicated his arguments²¹⁰. In terms of Sidgwick's philosophy, Parfit states quite simply that *ME* is 'the best book on ethics ever written'. Where Sidgwick had remarked that he had "two masters: Kant, and Mill", Parfit states his two masters to be Sidgwick and Kant, and *ME*, for Parfit, contains the 'largest number of true and important claims'²¹¹. He once said in a letter to Marcus Singer – "if there is any book which it would be for the best that everyone doing moral philosophy could assume that everyone thoroughly knew, it would be *ME*"²¹². What we have in Parfit's account of Sidgwick is - at both a personal and an academic level – the most important modernisation of Henry Sidgwick to date²¹³.

In terms of Parfit's own moral theory, it has already been stated in the introduction that there is neither room nor scope in this thesis to engage with it as fully as it deserves. It is, however, quite possible to identify where Sidgwick is directly relevant to Parfit's own influential arguments, and from this we can take several important points regarding one of the ways in which Sidgwick is presented today. On some of these points, my reading of Sidgwick aligns closely with Parfit's.

²⁰⁹ Both Parfit's 1984 book *Reasons and Persons* (Oxford: Clarendon Press: 1984), and *On What Matters* (Vols. 1 and 2) (Oxford: Oxford University Press: 2011) are held up by Bart Schultz as being 'almost wholly responsible for defining the 'field of analytical rehabilitations of Sidgwick's arguments' (Schultz, 'Review of Singer and Lazari-Radek, *Point of View of the Universe: Sidgwick and Contemporary Ethics*', University of Notre Dame, 7/7/2014).

²¹⁰ Parfit does admit that Sidgwick's work is not exactly easily digestible, and that Sidgwick is sometimes boring. Parfit also draws attention to Sidgwick's own description of his own "one damning defect of long-windedness and difficult dullness" – but on this point Parfit disagrees with Sidgwick, arguing that what Sidgwick thought would appear 'dull and repellent' is instead rather 'finely expressed' (Parfit, *On What Matters*, Vol. 1, p.xxxvii). I could not agree more.

²¹¹ Ibid, p.xxiii

²¹² See M. Singer (ed.) *Essays on Ethics and Method*, p.xii. Parfit also points out that one of Sidgwick's greatest merits was the accommodating way in which he would respond to his critics (which infuriated them), and that whereas he might not have been as *original* a philosopher as, say, Kant, Sidgwick is more like Darwin in that he took ideas that were clear, and grounded in common sense, and – to use Parfit's own perfect paraphrasing – "intensified them almost to the point of genius" (*On What Matters*, Vol. 1, p. xl).

²¹³ Parfit also argues that it is important we read the preface to the sixth edition (*On What Matters*, Vol. 1, p.452).

In his first volume, the critically acclaimed *Reasons and Persons* (1984), Parfit's central question is quite simply, what do we have reason to *do*? What, ultimately, are our reasons for acting? The entire first half of *Reasons and Persons* is concerned with this search, and of course Parfit shares this aim to unite ethics and rationality with Sidgwick²¹⁴. Echoing Sidgwick's introduction to *ME*, Parfit outlines three bases for what we have reason to do – self-interest (the most supremely rational aim for each human is that life goes best, or that an action is only rational if it maximises one's own interest), consequentialism (the one ultimate moral aim is that outcomes be as good as possible), and common sense morality (the belief that there are certain moral aims that each of us ought to achieve²¹⁵). In *On What Matters*, Parfit devotes an entire chapter to Sidgwick's own observation of the apparent stalemate between these methods that seems to exist in human rationality²¹⁶. Parfit's own

²¹⁴ The second half of *Reasons and Persons* is concerned with the distinctiveness of persons, and what makes us particular people. In this area too, Parfit draws on Sidgwick. Parfit frequently references Sidgwick's emphasis on the separateness of persons, defining it as a deep truth about personal identity, and as being one of the "fundamental facts underlying all reasons for acting" (*Reasons and Persons*, pp.133, 329. See also *On What Matters*, Vol. 1, p.133). The fact that we are all different people, with different lives to lead, is, according to Parfit, of great moral and rational significance. From this, it appears to be fairly clear that Parfit believes Sidgwick to be a forerunner of Rawls' separateness of persons theory. This puts him at serious odds with Peter Singer, who does not believe that Sidgwick suggested the separateness of persons as Rawls' interpreted it.

²¹⁵ *Reasons and Persons*, p.99. Parfit draws a direct correlation between his self-interest theory, and Sidgwick's theory of rational egoism. All of these theories are what Parfit calls, 'self-defeating' as rational methods. Self interest theory is a self-defeating argument because it often indicates that it would actually be more rational for us to believe some other theory, or act in some self-denying way, because *that* is actually what will make things go better for us. Consequentialism self-defeats because if we were all to become pure do-gooders, we would not be acting on our motives of desire (such as love for our families and the desire to work well etc.) and thus happiness would not actually be increased, but instead greatly reduced. Common sense morality too fails, on the basis that our prioritising of our own duties, ahead of the enabling of others to carry out their duties, may easily result in nobody being able to carry out many of their duties.

²¹⁶ This is of course what is represented by Sidgwick's Dualism of Practical Reason. It is this conflict between rational impartialism and rational egoism that so concerned Sidgwick that concerns Parfit here. After much discussion however, Parfit eventually refutes Sidgwick's DoPR view in as little as three sentences. It is simply impossible, Parfit argues, to have *most reason to act in two different ways*. If duty and self-interest appear to conflict, we simply have to revise or refute one of the views. Parfit's choice is to revise self-interest, which in turn refutes the problem by showing how self-interest theory itself makes it rational to act morally even if we know this would be worse for us. The details of this are given in *Reasons and Persons* (pp.19-21). During the discussion of whether self-interest can be overridden by morality when the two conflict, Parfit draws out the argument that self-interest theory can in fact support the belief that it is rational to keep promises (promises being the example of moral behaviour), even when we know that this will be worse for us. This is due to the fact that if self-interest can make us have this belief (which we sometimes do), then this automatically shows that that belief is justified. And if this is the case then there are many cases in which self-interest theory itself tells us that it is *rational to act morally* even when we know it would be worse for us; that is, moral reasons supersede self-interest, without actually denying it. Is it possible to say then, that Parfit solves Sidgwick's infamous Dualism of Practical Reason? To some extent, the answer would seem to be yes. Parfit refutes self-interest theory with far more conviction than Sidgwick, but while also

approach to this apparent plurality of reasons for acting (which is also at the heart of my claim for moral ambivalence made in the thesis introduction), is also similar to Sidgwick's, in that he believes these theories to not work by themselves, and to be in need of revision. Again following Sidgwick's view, Parfit admits that we all accept some form of common sense morality, but that it is in need of revision. The revision that Parfit suggests is given all the theories seem to be rational to some extent, and are therefore difficult to refute entirely, it would seem sensible to consider the "very attractive notion", as Parfit calls it, of perhaps somehow binding them together. Even more specifically, Parfit proposes that the *failings of consequentialism* require revisions that are provided by *parts of common sense morality*, and that the *failings of common sense morality* require revisions that are *in part consequentialist*²¹⁷. To use Parfit's own term, he suggests that his arguments about these bases of reasons for acting may *dovetail*. "We might be able to develop a theory that includes a combines revised versions of both C and M", he says. "Call this, the *unified theory*"²¹⁸. This would of course eventually become the Triple Theory that he developed in *On What Matters*, and that was outlined for its relevance to my argument for Rational Benevolence in the thesis introduction. What is important here is that Parfit directly recognises Sidgwick's claim that he had unified utilitarianism and common sense morality. Parfit is probably the first philosopher to recognise this idea of a combination of principles in Sidgwick, in language that accords directly with that I have used in my own similar argument. But Parfit does not utilise Sidgwick's own 'unified theory', and there is a crucial reason for this. It is that Parfit is arguing for unification between common sense morality and *consequentialism*, not utilitarianism, whereas Sidgwick was arguing 'for a Hedonistic version of Utilitarianism'²¹⁹. Parfit takes the consequentialist line because this theory may apply to several principles about what makes outcomes bad (such as more inequality and deception, for example), and therefore it is already closer in nature to common-sense morality²²⁰. This is an interesting claim, both in terms of Parfit's own unified

accounting for its presence. Parfit does not deny that the conflict between rational egoism and rational benevolence is very real: He does, however, present a far more convincing case for acting morally when this seems to conflict with self-interest.

²¹⁷ *Reasons and Persons*, p.112

²¹⁸ *Ibid.*

²¹⁹ *Ibid.*

²²⁰ See *On What Matters, Vol. 1*, pp.168-169. To distinguish between the two positions, Parfit believes consequentialism to be defined as 'whether our acts are right or wrong depends only on facts about how it would be best for things to go', and utilitarianism as 'things go best when they go in the way

theory, and as against Sidgwick's. It may well be that on this basis, Parfit's reconciliation between deontological and 'utilitarian', or consequentialist, principles is indeed superior to Sidgwick's, especially if it were to be shown that Sidgwick's version of utilitarianism is somehow more restrictive than Parfit's. But it is points like this at which we reach the limits of how far this thesis can effectively engage with Parfit here. I can only point to the fact that Sidgwick's own reconciliation as it appears in *Rational Benevolence* is not picked up by Parfit, and then to the possibility of further work in the area of comparing the two 'unified' theories²²¹. Having said this about Parfit's rejection of Sidgwick's particular utilitarianism however, *On What Matters* sees Parfit apply Sidgwick to numerous places during the building of Triple Theory. He recurses to Sidgwick's view that it is impossible to compare impartial and self-interested viewpoints– and disputes it, but only to an extent²²²; he suggests that Sidgwick's version of Impartial-Reason-Act-Consequentialism is actually seriously morally demanding²²³, and he also occupies a curious space between the view that Sidgwick is an act utilitarian and the view that he is a rule utilitarian, that he then deploys to support his own view that a version of rule utilitarianism is the correct one²²⁴.

There are then two major conclusions to be drawn regarding Parfit's extensive and progressive use of Sidgwick's work. The first is that Sidgwick steadfastly remains a

that would, on the whole, benefit people most, by giving them the greatest total sum of benefits minus burdens' (*On What Matters*, Vol. 1, p.373).

²²¹ Parfit does, however, directly refer to *Rational Benevolence* as Sidgwick's moral theory, and he is the only other writer to do so (*Reasons and Persons*, pp.139-140). However, his treatment of *Rational Benevolence* is very different to mine. Parfit defines *Rational Benevolence* as 'reason requires each person to aim for the greatest possible sum of happiness, impartially considered' (*Reasons and Persons*, p. 139). *Rational Benevolence* is, in Parfit's terms, a 'pure' theory, in that it is both agent and temporally neutral, and for this reason it poses a very real threat to self-interested reasons for acting. But Parfit actually does not believe Sidgwick himself to have extended the argument for *Rational Benevolence* far enough, given that Sidgwick thought *Rational Benevolence* could still be challenged, by an aspect of Prudence. The reason for this, Parfit argues, is that Sidgwick failed to adequately account for the importance of what Parfit calls 'Present aim theory' (*Reasons and Persons*, pp.140-142). See also Schultz, *Eye of the Universe*, p.216). However, Parfit does not seem to acknowledge that Sidgwick did actually finally conclude that *Rational Benevolence* supersedes *Rational Egoism* anyway. It is true that Sidgwick does not offer the length and depth of argument to support this claim that Parfit does, but he does at least allude to the idea in *ME* that morality has more claim to rational acceptance than self-interest. What is also important about Parfit's treatment of *Rational Benevolence* is of course that he does not recognise it to be the sort of reconciliatory theory that I argue it is.

²²² *On What Matters* Vol.1, pp.134-137

²²³ *Ibid*, pp.168-169

²²⁴ See *On What Matters* Vol.1 pp.251-252, 404. This point is especially interesting in light of the Sidgwick as an act/rule utilitarian debate outlined above. Singer and de Lazari-Radek directly challenge Parfit's non-act-utilitarian reading of Sidgwick here, in *Point of View of the Universe*. See below.

utilitarian on Parfit's view. Parfit does not discuss the non-consequentialist-duty-based/deontological aspects of Sidgwick's work, and where Kant appears in Triple Theory, his maxims are being reformulated in rule-consequentialist and contractualist terms. The second conclusion is that Parfit does not, in his own reconciliatory theory, directly use the reconciliation offered by Sidgwick. It would also appear, certainly from *Reasons and Persons*, that Parfit believes Sidgwick to have attempted a unified theory on a purely hedonistic basis, which I dispute here. However, Sidgwick is a constant presence in Parfit's building of Triple Theory, and can be seen to be instrumental to Parfit's arguments for what will eventually become a Kantian form of consequentialism – and that is telling enough in itself for how it is possible to read Sidgwick.

1.3. Perceptions and Oversights of Sidgwick's Reconciliation Between Deontological and Utilitarian Principles

In the previous two sections, I have been concerned to demonstrate two things. The first is that Sidgwick has unanimously been read as a utilitarian, since his own day and in the century that has followed. The second is that there have been some important movements towards recognising that Sidgwick's work offers something slightly different to standard utilitarianism - and even that he includes rather deontological concepts - but that in none of these cases the full significance of this has been pursued. For the next two stages in this particular chronicle of Sidgwick's work, I aim to examine the work of recent and current theorists who have either specifically recognised Sidgwick to present some kind of reconciliation between deontological and utilitarian principles of morality, or who are of the utmost importance to current Sidgwick studies but who have *not* fully recognised this aspect of his work²²⁵. In all these cases, I argue that Rational Benevolence has not been held up as the embodiment of that reconciliation, and that any recognition of Sidgwick's reconciliation has not been presented as a synthesis, as I argue it to be.

²²⁵ By including Parfit in the previous section, and not here in the section on current and recent theorists, I do not mean to exclude him from this category - *On What Matters* was only published in 2011. He was included in the above purely on the basis that *Reasons and Persons* appeared in 1984, and that that volume suggested as early as the 1980s that Parfit saw Sidgwick's theory as an example of the sort of unified theory that Parfit would later go on to develop into Triple Theory, in *On What Matters*.

The section, which still follows a rough chronology in order to remain in keeping with the ‘narrative’ style in which I have aimed to present this chapter so far, begins with Marcus Singer’s case against Sidgwick’s interpretation of common-sense morality. It then moves through, among others, Brad Hooker and Antony Skelton’s argument for and against a coherentist element in Sidgwick, and considers Mariko Okano-Nakuno’s argument against the apparent ‘discrepancies’ between Sidgwick and Kant. The chapter then examines Peter Singer’s influential use of Sidgwick, which in turn engages David Phillips, Thomas Hurka, and Roger Crisp. Although these last theorists are not always directly concerned with the idea of reconciliation per se in Sidgwick, it is my view that their work on Sidgwick is indispensable to any reading of him, and especially to readings of Sidgwick’s notion of *duty*. It is also still the case that these important readings of Sidgwick have not interpreted his theory of Rational Benevolence as I do here. Where it is with regret that there is not room to reference all theorists who have contributed to current readings of Sidgwick, as many as possible are included in the notes.

1.3.a. Marcus Singer: An Early Rejection of Sidgwick’s Reconciliation

Following the faint historical recognitions of Sidgwick’s nuanced version of utilitarianism, it might seem odd that the section on modern readings of Sidgwick should begin with an outright rejection of Sidgwick’s attempt to unite divergent moral principles. But Marcus Singer had initially utilised Sidgwick’s reference to Kant in his own argument for the ‘generalisation principle’ in his 1955 volume *Generalisation in Ethics*, and this generated for Singer both a lifelong fascination with Sidgwick, and an eventual disagreement with the idea that Sidgwick had achieved unity between common sense morality and utilitarianism²²⁶. Firstly, Singer believed Sidgwick to

²²⁶ Singer, *Generalisation in Ethics* (New York: Alfred A. Knopf: 1955). When Singer’s argument began life in 1948, it was called ‘Generalisation in Sidgwick’s Ethics’, but Singer admitted that this was not a good paper, and the revised 1955 version replaced it completely. The later version does still include Singer’s original appeals to Sidgwick’s arguments. Singer’s generalisation argument presupposes what Singer calls ‘the generalisation principle’, which is ‘what is right (or wrong) or one person must be right (or wrong) for any similar person in any similar circumstances. The parallel with Kant’s universalizability principle is clear, but notably it is actually Sidgwick to whom Singer immediately refers for support, quoting Sidgwick’s principle of justice (or impartiality), which states that ‘we cannot judge an action to be right for A and wrong for B, unless we can find in the natures or circumstances of the two some difference which we can regard as a reasonable ground for differences in their duties’ (III:XIII:380). The argument was not universally well received, with many

have broken his own rule in *ME*, by first claiming that he was not intending to introduce or endorse *the* method of ethics, and then proceeding to claim that the morality of common sense was in need of systemisation, which he then provides via the principle of utilitarianism²²⁷. More seriously, Sidgwick had, on Singer's view, misinterpreted the nature of common sense morality. In his 1986 article 'Common Sense and Paradox in Sidgwick's Ethics', Sidgwick's connection between common sense morality and utilitarianism shows up as a major concern for Singer²²⁸. Singer points out that, for Sidgwick, common sense morality is a body of moral truth, and that it is identified with intuitionism, so much so that he slides back and forth between them²²⁹. But for Singer, this is a mistake, stating the reason as 'Intuitionism is a philosophical theory, common sense is not'²³⁰. Singer believes Sidgwick to be wrong in declaring common sense morality unable to answer some questions in some situations, and wrong in his claim that common sense morality is often divided against itself: It is, therefore, wrong for Sidgwick to have 'foisted' utilitarianism on

writers and theorists disputing its validity, and the line of reasoning (see David Keyt, 'Singer's Generalisation Argument', in, *Philosophical Review* Vol.72 (1963) 466-476; A. Phillips Griffiths, 'Review of *Generalisation in Ethics*', in, *Philosophical Books*, Vol.3, No.1 (1962) 18-21; Howard Sobel, 'Generalisation Arguments', in, *Theoria* Vol.31, No.1 (1965) 32-60; Lansing Pollock, 'A Dilemma for Singer', in, *Philosophical Studies*, Vol.33, No.4 (1978), 425). But it has gained a lasting place in the ranks of arguments that attempt to provide a rational foundation for morality, and Sidgwick's input is undeniable. Singer's 1974 paper 'The Many Methods of Sidgwick's Ethics' (in, *The Monist*, Vol.58, No.3 (1974) 420-448) gave a careful and detailed biographical account of the birth and revisions of *ME*, and assessed Broad's claim that *ME* was the best treatise on ethics ever written (and, by proxy, Brand Blanshard's agreement with Broad that appeared in Blanshard's *Reason and Goodness*). But he also quotes Arthur Murphy's critique of *ME*, given in Murphy's *The Theory of Practical Reason* (LaSalle, Ill: Open Court: 1964), which stands in the sharpest contrast to Broad's opinion (see Singer, 'The Many Methods', pp.430-431), and cites the Dualism of Practical Reason as a reason why Sidgwick's work should not, perhaps, be considered to be the best treaty on morality ever written. Singer's argument here is convincing: 'Sidgwick regarded the first principle of rational egoism as self-evident. He also regarded as self-evident the first principle of utilitarianism, the axiom of 'rational benevolence' ('The Many Methods', pp. 445). It was the contradiction, or apparent contradiction between these that constituted for him, 'the profoundest problem of ethics', the Dualism of Practical Reason. Yet, strange as it may seem, he did not notice that his own conditions for self-evidence are violated by just this contradiction' ('The Many Methods', p. 446) Singer also points out that James Seth made this same observation in 1901.

²²⁷ 'The Many Methods', p.441. This point of Singer's provides an interesting parallel with my claim that Sidgwick did more than simply expound the various methods that we *do* use, and actually developed something close to a suggestion as to what we *ought* to do. Singer's reasoning as to why the correlation between intuitionism and utilitarianism amounts to a 'genuine method of ethics' (p. 442) is extremely sound, and correlates closely with the arguments for Rational Benevolence I will develop later. As Singer puts it: "despite his disclaimers in the introduction, Sidgwick wanted his *Methods of Ethics* to have practical bearings, and thought that it did" (p. 445).

²²⁸ In, *History of Philosophy Quarterly* Vol.3, No.1 (1986) 65-78.

²²⁹ Ibid p.66. See also Singer 'Sidgwick and Nineteenth-Century British Ethical Thought', in Schultz (ed.) *Essays on Henry Sidgwick*, pp.65-92.

²³⁰ 'Common Sense', p.69.

common sense, as a means of answering those questions²³¹. Singer quite clearly recognises that Sidgwick believed there to be little distance between Utilitarianism and Intuitionism – as is seen in his introduction to the 2000 collection of Sidgwick’s writings, *Essays on Ethics and Method* where he states that ‘the Utilitarianism he ends up with is Utilitarianism on an Intuitionistic basis, and Utilitarianism, by parity of reasoning, can be said to be unconsciously intuitionistic’²³². But for Singer, as his 1986 paper makes clear, this reconciliation is frustrated by the fact that Sidgwick has fundamentally mishandled the common sense morality that is represented by that Dogmatic Intuitionism, by equating the two²³³. This is, I believe, quite a unique charge in the world of objections to Sidgwick. But it also depends on our taking seriously Singer’s distinction between “intuitionism” as an epistemological method of ethics (that which I distinguish with a small ‘i’) and ‘Intuitionism’ as representative of Sidgwick’s view of common sense morality, and thus the deontological position (the type I distinguish with a capital ‘I’). As I outlined in the thesis introduction, and will argue in full in the next chapter, it is my view that Sidgwick’s intuitionism and his Intuitionism are united - and therefore it is not problematic when that common sense morality is united with the intuitions that lead to the utilitarian principle²³⁴.

²³¹ Ibid. p.73

²³² Singer (ed.) *Ethics and Method*, p.xxi

²³³ Singer remarks in *Common Sense and Paradox* that Sidgwick took Rational Benevolence to be a self-evident principle, and that he considered it to be the basis for the utilitarian principle (p.75). Singer thus confirms that whereas he might not agree with the link between utilitarianism and common sense morality, this is at least how Sidgwick saw it. Singer also states here that he himself is not prepared to accept rational benevolence as a principle of the morality of common sense (p. 75). Nor, however, incidentally, is he prepared to accept Sidgwick’s account of rational egoism as self-evident.

²³⁴ Marcus Singer has one more important contribution to make to the Sidgwick cause – in this case, a historical, rather than a philosophical one. This is Singer’s account of his personal engagement with Sidgwick that we are given in the preface to *Ethics and Methods*, and its illumination of the events that occurred in the 1980s, which concerns the general raising of awareness of Sidgwick’s work. Singer describes how during the 1960s, when he first thought about compiling the material that was eventually to become his volume on Sidgwick’s essays and methods, there simply was not enough interest in Sidgwick to warrant the project. Papers such as Singer’s own of 1974 were at that time relatively rare. In 1983 however, Singer established The Sidgwick Society, the purpose of which was, according to its mission statement, simply ‘admiration for the character of Henry Sidgwick and the way he practiced philosophical inquiry: careful, cautious contemplation, considering all points and every possibility, along with his resolute aim at truth and clarity and clarified understanding.....and his resolute good sense, good judgements, and good will’ (see *Ethics and Method*, p.x). Singer also quotes from a later statement, which is deserving of recognition here: ‘The society does not do anything – it holds no meetings, publishes no journals....it just *is*. It exists as a tribute to Henry Sidgwick, and the example he set in his intellectual work and life’. Surprisingly, for both Singer and the reader, Singer recalls his amazement at the volume of interest that was expressed in the society - at how many people wanted to become members. Interest in Sidgwick, he realised, was far greater than he or even his partial collaborator Bart Schultz had ever thought. Academics from Richard Brandt to William Frankena have at one point been on the Board of Directors of The Sidgwick Society,

1.3.b. Janice Daurio: *The Methods as One Single Moral Theory*

Among the works that most specifically read Sidgwick in terms of how far his work provides a kind of unity between moral principles, Janice Daurio's 1997 paper 'Sidgwick on Moral Theories and Common Sense Morality' is notable, mainly for the extremity of its claim²³⁵. The paper is in part a direct challenge to Schneewind's claim that one of the reasons for the historical neglect of ME is its lack of unified argument²³⁶, and in part her own response to her observation that 'philosophers since Broad have continued to miss the essential unity'²³⁷. On this point about unity having been missed in *ME*, I am obviously in direct agreement with Dario. Where I do not agree, however, is with Daurio's argument is that the entirety of *ME* is actually 'a thorough and systematic examination of a single moral theory', one which unifies utilitarianism, intuitionism, *and* egoism²³⁸. Clearly, this goes even beyond my argument for Sidgwick's reconciliation. But Daurio's theory is, like M. Singer's, helpful for illuminating what I am *not* saying about the reconciliation for which I am arguing.

Daurio makes the argument for the single moral theory from the apparent inconsistency in *ME* between Book III, in which Sidgwick agrees with common sense moralists who do not believe that a notion of goodness is needed in order to establish what is right conduct, and the end of Book IV, where Sidgwick argues that rightness wholly depends on the universal happiness. This, Daurio admits, gives a contradictory impression, but can be overcome if all three of the moral theories Sidgwick set out to investigate – egoism, intuitionism and utilitarianism – are part of "that single moral theory...the one implicit in common sense morality"²³⁹. Her

and as Singer was writing in 2000, it had just opened its website, which is still active and running today.

²³⁵ In, *History of Philosophy Quarterly* Vol.14, No.4 (1997) 425-445. Daurio claims in the introduction to this paper that Sidgwick work is 'still in its infancy', even at this relatively late year. Neatly following the section above that outlined Marcus Singer's establishing of The Sidgwick Society, it was actually Singer himself who picked Daurio up on this statement, politely pointing out in a private letter to Daurio that Sidgwick studies would be better considered to be 'at the kindergarten stage' ('Sidgwick on Moral Theories', p.442).

²³⁶ See Schneewind, *Sidgwickian Ethics*, p.72

²³⁷ 'Sidgwick on Moral Theories', p.426

²³⁸ Ibid. p.426.

²³⁹ Ibid., p.427. Obviously, this claim flies directly against Sidgwick's statement that he was *not* aiming to defend one theory above all others – and nowhere does Daurio show that she is aware of this statement often repeated by the author. Sidgwick did not even claim to be defending *one* of the three methods, let alone a situation in which all three can be viewed together.

argument is based on an understanding of a 'moral theory' as a complex combination of methods and principles. The relationship of the rightness of the rules to the good of the outcome is Daurio's interpretation of what Sidgwick meant by a method: Then the 'good' itself is a principle in this system. Central to Daurio's argument is that Sidgwick believed various aims or methods can be paired with various ends or principles. That is, within one single theory, a method might be attached to some principle, or more than one method might be attached to a principle. This leads her to declare that 'the moral theory implicit in common sense morality – and therefore the moral theory of *The Methods* – is this complex moral theory, linking the three methods of egoism, intuitionism, and utilitarian, to some end. If it can be shown that there is just one ultimate end, and if it can be shown that this one ultimate end is the one achieved by agents acting on rules generated by all three methods, then 'common sense morality is justified'²⁴⁰. Daurio uses a quote from Sidgwick that appears to support her argument that 'the good' unifies all of the moral theories into one. Sidgwick asks that if universal good is not what systemises human activities, then what can it be? If this systemisation of activity can be achieved, Daurio says, then the three methods naturally become linked by their mutual reference to Happiness as the ultimate end – all three methods in common sense morality are linked to just one principle, the utilitarian one.

In some ways, it is easy to see why Daurio theorises *ME* to be a coherent whole, with common sense morality emerging as the overall moral theory, providing a sort of playing field in which all the various 'methods' (egoism, intuitionism and utilitarianism) can sub in and out in a pursuit of the goal of the good. But Daurio's paper continually runs a high risk of making the unity at which she is aiming feel rather forced. There is, for example, a fairly significant and inexplicable leap between the claim that 'the discovery of three methods in common sense morality means that the moral theory implicit in common sense morality is complex' and the sentence immediately succeeding that one in which it is stated that 'if the three methods of egoism, intuitionism and utilitarianism are successfully joined to just one principle, whatever that principle is, then common sense morality is justified: It presupposes a valid, complex moral theory'. This is a very sweeping statement to make, when there

²⁴⁰ Ibid. p. 434

is little material evidence of such a principle – or even little reason for there to be a particular principle at all ('whatever that principle is')²⁴¹.

But ultimately it is the inclusion of all three methods into one systemised morality that aims at Happiness that is the most problematic aspect of Daurio's theory. At the mid-way stage of the paper, Daurio points out that Intuitionism, for Sidgwick, has led directly to utilitarianism²⁴². Daurio comes tantalisingly close here to recognising that intuitionism and utilitarianism are intrinsically linked, but she does not focus on the *nature* of the link between them, instead moving straight on to argue that common sense morality also implies egoism, which has the abrupt effect of bringing Daurio's theory to a sudden, frustrated close: "The sad conclusion: We are stuck with the egoist method, which cannot be logically or plausibly connected to the utilitarian principle in the indirect moral theory of common sense morality"²⁴³. Daurio's efforts have, essentially, been thwarted by that most ancient of *The Methods*' problems – the Dualism of Practical Reason. This, she argues, is why *ME* cannot be read as a unified argument – but she does not take this to mean that unity is not present in *ME* at all.

On this last count, I do not disagree with Daurio. Unity does exist in *ME*, and it does provide a moral theory. I am also in complete agreement with her observation that unity in *ME* has been missed because so many theorists have simply regarded *ME* as a defence of *one* moral theory – utilitarianism. However, as Daurio's ultimately

²⁴¹ Ibid. p.433. Daurio has, earlier in the paper, shaped what is meant by this principle to some extent, for example stating that the connection between a theory's aim and the end achieved is meant to be logical, and that if this can be proven then the moral theory is valid (p. 430). However, such arguments are effectively denounced by Peter Singer's refuting of Rawls' argument that Sidgwick used reflective equilibrium, on the basis that Sidgwick did not believe that just because a moral theory could be seen to be valid, does *not* necessarily make it right. Daurio's argument for the connections between the various theories also consists in arguing that some moral theories are direct, by which it is meant that the conscious aim of the method and the ultimate end of the principle are the same i.e. egoism generates rules that are to do with the individual's happiness as the ultimate end, and acting on those rules leads to said happiness, and indirect theories in which the conscious aim is matched with a principle of another sort i.e. where egoism generates rules that are to do with the agent's happiness, but acting on those rules might actually defeat that very aim (pp. 432-434). These claims are not wrong, and Daurio is absolutely correct to point out that Sidgwick anticipated some of the later refinements which utilitarianism was to undergo, in the hands of Derek Parfit for example, and Dean Cocking and Justin Oakley in their defence of indirect consequentialism (Cocking and Oakley, 'Indirect Consequentialism, Friendship, and the Problem of Alienation, in, *Ethics*, Vol.106, No.1 (1995) 86-111). But in the case of *ME*, it seems too far a leap to use this mechanism as a means of arguing for a broad-spectrum unity across all of the methods. Sidgwick may have perceived that methods and ends are often indirectly, rather than directly, related, but it does not follow that his investigation of mankind's various ways of making moral decisions itself actually uses this in practice.

²⁴² 'Sidgwick on Moral Theories', p.440

²⁴³ Ibid.

foiled argument shows, reconciling all three methods is not where that unity exists. It is simply not the case that *ME* overall is one single moral theory. Sidgwick's moral theory depends on the reconciliation of two of the methods only. Daurio was not wrong to look for moral unity. But she was looking in the wrong place.

1.3.c. Brad Hooker: Synthesis as Coherentism in Sidgwick's Work

In the past decade or so, the highly interesting question has arisen of whether - in the propositions, claims, and conclusions that Sidgwick puts forward – Sidgwick is advocating a coherentism theory of knowledge, or a foundationalist one. There are very good grounds in *ME* for holding either position, and this particular debate also gives rise to some important points about the nature of Sidgwick's reconciliation between deontological and utilitarian principles as it appears in *Rational Benevolence*. Whereas there are several theorists who contribute importantly to this topic, I have taken Brad Hooker and Antony Skelton to present the most direct discussions; they are also the most relevant for the fact that they connect their coherentist/foundationalist investigations in some way to an idea of the relationship that Sidgwick developed between deontological and utilitarian principles – as I will also do.

Hooker's 2000 paper 'Sidgwick and Common Sense Morality' begins with a paragraph overflowing with praise for *ME*, which includes Hooker's belief that Sidgwick's criticisms of Kant were 'exactly right', and that his account of moral judgement is better than any other²⁴⁴. For Hooker, Sidgwick's main contributions to moral philosophy are his moral epistemology (which he identifies as having been considered foundationalist), his use of common sense in utilitarianism, and his development of utilitarianism itself. From this then it is clear that Hooker joins the ranks of those who believe Sidgwick to be a utilitarian. But Hooker does recognise that Sidgwick was 'clearly looking for a moral system that would actually resolve uncertainty and disagreement'²⁴⁵. More specifically, Hooker emphasises that Sidgwick believed common sense morality to be in need of – and he uses Sidgwick's phrase here - *rational synthesis*²⁴⁶. That is, Hooker disputes the view that Sidgwick's epistemology was wholly foundationalist. Hooker outlines Sidgwick's criteria for self-

²⁴⁴ Hooker, 'Sidgwick and Common Sense Morality', in, *Utilitas*, Vol.12, No.3 (2000) 347-360.

²⁴⁵ Ibid. p.354-355.

²⁴⁶ Ibid. p.351

evidence, and argues two things; firstly, that the principles of common sense morality that Sidgwick passed through these criteria failed, thus challenging the foundationalist reading, and secondly that as Sidgwick then reintroduces some common sense morality principles in his argument for utilitarianism, this produces a coherentist system²⁴⁷. In this system, on Hooker's view, it is utilitarianism that provides the necessary systemisation, reforms, and revisions²⁴⁸.

The theory of coherentism posits that beliefs within a system must cohere with one another in order to be valid, or justified²⁴⁹. Hooker's point that the rational synthesis of our moral beliefs is, in Sidgwick, a form of coherentism is a good one, and it does seem as though this might explain the relationship between common sense morality and utilitarianism. I do not disagree with Hooker: In fact, Hooker's account of the need for this relationship accords closely with the one I will give in the next chapter, in the discussion of how far Sidgwick believes common sense morality to be valuable by itself²⁵⁰. And given that definition, coherentism would appear to be a valid alternative means of representing the idea of the reconciliation between deontological and utilitarian principles in Sidgwick's work: It supports fully my claim outlined in the introduction (and expanded later in chapter 4) that both the deontological and utilitarian properties of Sidgwick's moral theory can *only* be

²⁴⁷ Hooker also points to Sidgwick's observation that our moral intuitions (those that comprise common sense morality) are often vague, and loose, and that this does not accord with a 'good intuition'; "In order for an intuition to be a good one, careful reflection should not drain our confidence in it. But that is exactly what Sidgwick thinks happens with many of the intuitions that make up common sense' (p. 350).

²⁴⁸ As Hooker says, 'Utilitarianism makes sense of common sense morality, by endorsing most of it, and explaining its unity' (p. 351). In this paper, Hooker also furthers the 'Sidgwick as an act/rule utilitarian' debate. Hooker admits that there are various ways in which Sidgwick could be interpreted in these areas, but he is himself mainly of the opinion that 'indirect collective utilitarianism', or a form of rule utilitarianism, would best represent Sidgwick's position (p. 360), and that act utilitarianism is probably not the version that coheres common sense morality in order to create the rational synthesis.

²⁴⁹ The type of coherentism in question here is the coherence theory of epistemic *justification*, as opposed to the coherence theory of truth. Other notable theorists to debate Sidgwick's coherentist epistemic justification are Bart Schultz, David Brink, J.B. Schneewind, Robert Shaver, and Roger Crisp (see below). For Hooker's further discussions of coherence, rule consequentialism, and reflective equilibrium, see Hooker's *Ideal Code, Real World*, p.9; 'Rule-Consequentialism', in, *Mind*, Vol.99, No.393 (1990) 67-77; 'Rule-consequentialism, Incoherence, and Fairness' *Proceedings of the Aristotelian Society* Vol. 95 (1995) 19-35; 'Ross-Style Pluralism versus Rule-Consequentialism' *Mind*, Vol.105, No.420 (1996) 531-552; Hooker, Elinor Mason and Dale E. Miller (eds.) *Reflective Equilibrium and Rule Consequentialism*, in *Morality, Rules and Consequences* (Edinburgh: Edinburgh University Press Ltd: 2000). See also Dale Miller, 'Hooker's Use and Abuse of Reflective Equilibrium', in Hooker et al (eds.) pp.156-178.

²⁵⁰ Incidentally, Hooker also takes common sense morality to refer specifically to duties (and virtues), as I do. 'Sidgwick and Common Sense Morality', p.360.

validated, or justified, when they are brought into harmonious relationship with each other.

However, I do not believe this is the end of the coherentist/foundationalist part of Sidgwick's story. Firstly, there is the issue of Sidgwick's epistemic intuitionism. Sidgwick quite clearly defends a highly non-inferential form of intuitionism (we see this particularly in his statement in the PD that he "must somehow see that it was right....I was forced to recognise the need of a fundamental ethical intuition' (xvi), italics Sidgwick's), and this search for a non-inferential knowledge continues in Sidgwick's search for ultimately rational self-evident first principles²⁵¹. All of this strongly indicates foundationalism. Then the coherentist aspect appears again when Sidgwick *derives* the point of view of the universe, and thus Rational Benevolence, from the self-evident Kantian maxim, presented by Sidgwick as Justice. Rational Benevolence thus appears non-foundationalist, as it is inferentially known. Sidgwick's argument that Rational Benevolence is the *proof* of utilitarianism is also clearly non-foundationalist. At a wider level, this then translates to a relationship between utilitarianism and common sense morality in which the two revise and support each other – and this would seem to agree with the coherentism that Hooker assigns to Sidgwick. It also accords with the way I have expressed Rational Benevolence, which is that the rationality and the justification of each principle depends directly on the rationality and validity of the other; knowledge is clearly inferential here, and *also* a form of coherentism²⁵².

²⁵¹ It is on the basis of this foundationalist intuitionism that Sidgwick and W.D. Ross are often united.

²⁵² Roger Crisp ('Sidgwick and the Boundaries of Intuitionism', in, Phillip Stratton-Lake (ed.) *Ethical Intuitionism* (Oxford: Clarendon Press: 2002) 56-75) and David Brink ('Common Sense and First Principles in Sidgwick's Methods' in, *Social Philosophy and Policy*, Vol.11, No.1 (1994) 179-201) both contribute importantly to this debate, both using as their starting point Sidgwick's epistemological intuitionism, and arriving at different conclusions as to just how effective this type of knowledge is. These conclusions, it can be seen, seem to depend on whether one is willing to read Sidgwick as a foundationalist, or a coherentist who was actually using – whether he intended it or not – a form of reflective equilibrium. Crisp, who has taken the foundationalist approach, has greater faith in Sidgwick's intuitionism. He describes intuitionism as being for Sidgwick 'a doxastic faculty', with a 'capacity to form non-inferential, self-evident beliefs that certain actions, rules, or whatever are right and reasonable, and moral intuitions are such beliefs' ('Sidgwick and the Boundaries of Intuitionism' p.57. See also Crisp, *The Cosmos of Duty*, p. 97). Crisp believes that this Sidgwickian version of intuitionism is by far the strongest version of that method of knowledge. Brink, on the other hand, picks up on the inferential process of moral justification used by Sidgwick in the search for self-evident principles, and the fact that Sidgwick clearly believes this process to be central to the establishment of those principles. This process, Brink argues, in which principles are 'discursively justified' is a form of dialectical equilibrium ('Common Sense and First Principles', p.192) – and this is inconsistent with the sort of epistemic intuitionism Sidgwick also advocated. Russell Hardin is also of the firm position that a form of reflective equilibrium is present in Sidgwick's work ('Common Sense at the Foundations', in,

The question becomes then, is the ‘synthesis’ between moral principles – the one that Hooker refers to here and the one that I argue is present in Rational Benevolence – a coherentist one, as Hooker would have it? And might this be a credible alternative way of understanding the relationship between deontological and utilitarian principles? It is my view that Sidgwick cannot be defined as either a coherentist or a foundationalist. I believe that there is a peculiar mixture of coherentist and foundationalist elements in Sidgwick’s work, and that this is the direct result of his peculiar mixture of Reason and intuitionism, and that – also like his Reason and intuitionism – they seem to be inherently connected. His reasoning establishes a non-inferential, foundationalist basis, and this *results* in a form of coherentism. Rational Benevolence is therefore, I argue, a foundationalist/coherentist hybrid²⁵³. I do not, unlike Brink, think that this causes any damage to Sidgwick’s epistemology, or his moral theory. For one thing, it is not absolutely necessary to define Sidgwick either way; it is not important per se to how we understand Rational Benevolence that it is coherentist or foundationalist²⁵⁴. But it does serve to show some of the character of Rational Benevolence, and of Sidgwick’s individuality as a philosopher.

Schultz (ed.) *Essays on Henry Sidgwick*, pp.143-161). The fact that Sidgwick appears to be using both an intuitionist and a dialectical process of justification does indeed seem to point straight to reflective equilibrium, which is a form of coherentism, and would seem to accord with Hooker’s view of Sidgwick. Hooker himself argues that Sidgwick did advocate a form of reflective equilibrium (*Ideal Code, Real World*, p.9). The reflective equilibrium/coherentism/foundationalism debate continues, as will be seen below in the work of David Phillips, and Peter Singer.

²⁵³ David Phillips agrees specifically that Sidgwick’s epistemology is a hybrid (*Sidgwickian Ethics*, p. 61).

²⁵⁴ It could, however, alternatively be argued that it *is* actually important that we know whether Rational Benevolence is coherentist or foundationalist, as whether the principles within it are inferentially or non-inferentially known may affect my argument for the mutually informing relationship between deontological and utilitarian principles that is so vital to the synthesis. To this end, I venture that a purely foundationalist view of Rational Benevolence is possible, if *both* the absolute-principle based and the utilitarian based principles are argued to be independently self-evident, and I admit that this might be the only version that my account of Rational Benevolence allows. On the other hand, we may, in the end, be forced to call Rational Benevolence coherentism, because foundationalism precludes any inferential knowledge, where coherentism does not necessarily (i.e. it is possible to hold a non-inferentially known belief, and for this to then generate other beliefs that cohere with it). I still maintain however that a coherentist/foundationalist hybrid may be a valid representation of Rational Benevolence that does not threaten the need for each moral property to be independently valid. Rational Benevolence is foundationalist in that *we only know each moral property because we know the other*, and coherentist in that *they are only validated because they cohere*. This may, in fact, even add to the strength of Rational Benevolence.

1.3.d. Antony Skelton: The Rejection of Sidgwickian Coherentism, and the Question of Reconciliation

Antony Skelton has been writing prolifically on Sidgwick for over a decade²⁵⁵. His work shows great admiration for Sidgwick, while also challenging some of Sidgwick's most famous claims, such as the self-evidence of certain principles that Sidgwick thought to have met those conditions. Skelton is relevant to the wider issue of how (or indeed if, on Skelton's view) Sidgwick brought dogmatic intuitionism together with utilitarianism via philosophical intuitionism – but coherentism is also a theme in that discussion, and it seems suitable to discuss this aspect first, following Hooker's version.

In his 2010 paper *Henry Sidgwick's Moral Epistemology*²⁵⁶, Skelton begins with the statement that Rawls clearly considered Sidgwick to be a coherentist, given that Rawls thought Sidgwick to be using a system of reflective equilibrium, and that Peter Singer disagreed on the basis that he believed Sidgwick to be committed to a foundationalist intuitionism which relies on self-evident principles as justification²⁵⁷. Skelton perceives Singer to have moved away from this position somewhat, by sometimes arguing that Sidgwick supports his arguments with reference to common sense morality²⁵⁸. But for Skelton, Singer's first view was the correct one. Skelton believes Sidgwick's moral epistemology to be an inherently foundationalist one, that allows no 'evidentiary role' to common sense morality²⁵⁹. Utilising Sidgwick's justification of utilitarianism as the model of Sidgwick's epistemology, Skelton emphasises Sidgwick's non-derivatively justified philosophical intuitions - that is, his criteria for self-evidence²⁶⁰. His careful selection of texts from *ME* lead him to the point at which he states that the 'rational synthesis' (i.e. the same rational synthesis from the same passage quoted by Hooker) is provided by those philosophical

²⁵⁵ See, as a selection, 'Utilitarian Practical Ethics: Sidgwick and Singer' (in, Placido Bucolo, Roger Crisp, Bart Schultz (eds.) *Henry Sidgwick: Ethics, Psychics, and Politics* (Catania: University of Catania Press: 2011); 'Henry Sidgwick's Practical Ethics: A Defence', in, *Utilitas* Vol.8, No.3 (2006) 199-217; 'Schultz's Sidgwick', in, *Utilitas*, Vol.19, No.1, 91-103 (2007), as well as the articles I have directly referenced in the main text, and his engagement with David Phillips, which appears below.

²⁵⁶ In, *Journal of the History of Philosophy*, Vol.48, No.4 (2010) 491-519.

²⁵⁷ Something that I agree exists in Sidgwick.

²⁵⁸ Something else that I agree exists in Sidgwick.

²⁵⁹ 'Moral Epistemology', p.492.

²⁶⁰ Skelton says "An explanation of the structure of the case for utilitarianism and the justificatory architecture to which it appeals is the chief task of any account of Sidgwick's epistemology' (Ibid. p.493).

intuitions²⁶¹. This is a thoroughly foundationalist position, in which the philosophical intuitions “exhibit the truth” of the first principle of utilitarianism²⁶².

This foundationalist view appears throughout Skelton’s work²⁶³. We might now ask what consequences this reading, as opposed to the coherentist one, has for Skelton’s wider understanding of the relationship between philosophical intuitionism and utilitarianism.

Skelton’s foundationalist reading of Sidgwick certainly makes it harder to argue for a synthesis, as I have described it. On Skelton’s view, philosophical intuitionism is the justification for utilitarianism, and this is the end of the matter – and this does not amount to the sort of mutually informing relationship between deontological and utilitarian principles that I wish to develop in this thesis²⁶⁴. Furthermore, Skelton’s

²⁶¹ What is particularly interesting here is that Skelton uses many of the same passages from *ME* in his argument for foundationalism that are quoted by Hooker in his argument for coherentism. For example, ‘although common sense cannot offer us any “independent and self-evident rules of morality” (III:XI:360), it “may still be perfectly adequate to give practical guidance to common people in common circumstances” (III:XI:361)’, (see Skelton, *ibid.* p. 494). As will be seen, these passages will appear again, in my own arguments for the synthesis between deontological and utilitarian principles. Skelton also makes the point here that just because we might be able to get our moral views to cohere, this does not necessarily mean that they justified. This is a further argument against coherentism, and strongly reinforces Skelton’s argument for a foundationalist reading of Sidgwick.

²⁶² *Ibid.* p.496. See also Skelton’s later statement that ‘moral justification is therefore foundationalist in nature; moral theories are justified by appeal to philosophical intuitions’ (p.497).

²⁶³ See Skelton’s starting point in his 2008 paper ‘Sidgwick’s Philosophical Intuitions’ (in, *Ethics and Politics*, Vol.10, No.2 (2008) 185-209), which is Sidgwick’s view that the utilitarian method could not be made coherent and harmonious without a fundamental intuition, and that the only moral intuitions that can be found to be ultimately sound are only those generated by the principle of utilitarianism (p. 185). In a later paper, ‘Henry Sidgwick: 1838-1900’ (in, William Sweet (ed.), *Biographical Encyclopaedia of British Idealism* (London: Continuum: 2010) Skelton is again clear that whereas the self-evident axioms of philosophical intuitionism indicate utilitarianism, it is utilitarianism that ‘systemises and corrects common sense morality’. He follows this up by quoting Sidgwick’s statement that ‘utilitarianism is the final form into which Intuitionism seems to pass, when the demand for really self-evident first principles is rigorously pressed’ (III:XIII:388).

²⁶⁴ Skelton has, at one point, actually picked up the coherentist interpretation, arguing that it ‘gathers steam’ from Sidgwick’s argument that utilitarianism can explain, systemise and make consistent the varied and often disparate rules of common sense morality, which in turn then lends utilitarianism a certain degree of credibility (‘Henry Sidgwick’, in, J. Mander and A. P. F. Sell (eds.) *The Dictionary of Nineteenth Century British Philosophers* (London: Bloomsbury: 2002). But Skelton rejects this view. For one thing, he says, Sidgwick’s argument that utilitarianism seems to lend a certain validity to common sense morality is in serious question. If what he means is that the claims of common sense are given probative value, then this might support a coherentist reading. But if this is the case, then Sidgwick has to explain why certain claims from common sense morality are discarded while others are not, *without* appealing either to common sense axioms themselves, or to utilitarianism (which would essentially equate to a tautology). Also, Sidgwick’s search for self-evident proofs itself needs to be explained, given that if one is a coherentist regarding the types of moral principle, then presumably non-inferentially justified belief of intuitions is unnecessary. I argue here however that Skelton’s argument does not seem to follow. If a coherentist position is understood to hold that a belief is justified if it coheres with other beliefs in a system, there is nothing here to necessarily eliminate self-evidence as being a part of that system. The property of self-evidence may well be a necessary part of that particular system of epistemology. David Phillips, who is of the opinion that

view leads him to quote Sidgwick's passage, "the Intuitionist method rigorously applied yields as its final result the doctrine of pure Universalistic Hedonism, which is convenient to denote by the single word, 'Utilitarianism'" (III:XIV:406-407), and use this to argue that on this *foundationalist* basis, Sidgwick 'takes himself to be reconciling Utilitarianism with an intuitionist (rationalist) epistemology'²⁶⁵. Skelton does not specifically pursue the idea of this reconciliation, focusing instead on furthering his argument that Sidgwick justified utilitarianism as a first principle on the basis of philosophical intuitions, and arguing that this same process is occurring in Book IV, when Sidgwick aims to 'prove' (as a first principle) the principle of utilitarianism. Where Skelton does recognise that others have interpreted the common sense morality/utilitarianism link to be one of coherentism, he later states that this is implausible, on the basis that he believes Sidgwick to be clear that the first principle (utilitarianism) cannot be derived from other propositions²⁶⁶. This is admittedly a strong position (if Skelton is to be agreed with – see note. 128), and brings into view the chief problem that Skelton's argument for foundationalism causes to my argument for the mutually informing relationship between deontological and utilitarian principles. This is that common sense morality, or the dogmatic intuitionism which is its philosophical version, could be thought of as having little validity of its own. Even where Sidgwick accepts some of those maxims as valid,

Sidgwick is a moderate foundationalist, suggests that such a 'hybrid' view results if 'on the one hand, we think of foundationalism as the view that epistemic justification is entirely a matter of deriving propositions which stand in need of justification from self-evident premises...and if, on the other hand, we think of coherentism as the view that justification is wholly a matter of relations between propositions, and not at all a matter of their self-evidence' (*Sidgwickian Ethics*, p.61). Phillips goes on to outline what is essentially also this thesis' position on Sidgwick's foundationalism/coherentism: 'Sidgwick's actual view then shares with foundationalism the idea that self-evidence has an essential epistemic role. But it also features two characteristically coherentist ideas; the idea that apparently self-evident propositions are not immune to doubt or questioning, and the idea that my justification for believing apparently self-evident claims comes in part from their coherence with other apparently self-evident claims I believe...'

²⁶⁵ 'Moral Epistemology', p. 497.

²⁶⁶ Ibid. p.503. Skelton also challenges Schultz's coherentist account of *ME*, in his 2007 paper, *Schultz's Sidgwick*, and by default also that of Robert Shaver (*Rational Egoism: A Selective and Critical History* (Cambridge: Cambridge University Press: 1999). If Sidgwick really was developing a coherentist justification of utilitarianism here, Skelton argues, 'then he would be keen to show that the theory really was derived from the rules and judgements of common sense morality. But this he does not do'. In my argument for Rational Benevolence however, the utilitarian principle *does* emerge as derived very clearly from rules and judgments that are at least related to common sense morality (see chapter 4). This would seem to mean then that I cannot argue for coherentism in Sidgwick, on Skelton's understanding. But this comes down to the fact that Skelton and I read Sidgwick's 'proof' of utilitarianism in very different ways. It is not my aim in this thesis to ultimately decide whether or not Sidgwick was a coherentist, but the diversity in interpretations over this point suggest that there is scope for further engagement here.

Skelton argues, this ‘does not mean that they have warrant apart from the first principle’. And this would seem to frustrate, or fundamentally challenge, my view that the utilitarian principle itself relies on a notion of duty, represented by Intuitionism, for its own establishment and validity²⁶⁷. But there is actually a point of Skelton’s with which I directly agree, and which I deploy here as a means of arguing that utilitarianism *can* be brought into a mutually informing relationship with deontological principles. This is his statement that Sidgwick did not, as coherentists want him to do, argue that utilitarianism is directly derived from the rules and judgements of common sense morality²⁶⁸. I also do not think this is how Sidgwick derives the utilitarian principle; it is not the rules and judgements of common sense morality that themselves point to utilitarianism, but rather the deontological properties that inform that position (which Sidgwick only finds to be truly self-evident and defensible in the form of Justice, and Prudence, and then Universal Benevolence). I admit that in Book IV Sidgwick is clearly showing how utilitarianism systemises and corrects common sense morality – but I maintain (and I will argue in full in chapters 3 and 4) that this is not where the relationship lies, as Skelton perceives the coherentists to argue²⁶⁹. I will shift the argument for a type of coherentism (as I said above, a foundationalist/coherentist hybrid) instead to Sidgwick’s argument for the necessary connection between a specifically deontological notion of moral obligation, and the utilitarian principle.

There is one final remark to be made about Skelton and his role in modern day readings of Sidgwick, which is that Skelton never disputes that Sidgwick is a utilitarian²⁷⁰. Skelton states in both the 2002 and the 2008 papers his view that

²⁶⁷ On the other hand, I am not arguing that utilitarianism depends directly on principles of common sense morality per se – this is also not what Sidgwick argued. I am arguing that utilitarianism is derived from self-evident, non-consequentialist-based principles (Sidgwick’s philosophical intuitions), and that whereas Sidgwick draws these out of common sense morality, they become something much different under his treatment to the other principles of that position, which Sidgwick finds to not be self-evident. It is these two truly self-evident principles (Justice and Prudence) which provide the grounds for the reconciliation between deontological and utilitarian moral properties.

²⁶⁸ ‘Moral Epistemology’, p.503

²⁶⁹ Nor is utilitarianism, however, adequate by itself for a system of ethics – and this fact is immediately represented by the meta-ethical connection between deontological and utilitarian properties embodied by Rational Benevolence.

²⁷⁰ Skelton actually gives an extremely succinct account of Sidgwick’s argument for utilitarianism, in 14 steps, in ‘Sidgwick’s Philosophical Intuitions’ (pp. 203-204). I do not dispute Skelton’s presentation of the process through which Sidgwick arrived at a utilitarian position – I do, however, dispute that that utilitarianism is *purely* utilitarian. Having said this, it would be unfair not to recognise where Skelton perceives Sidgwick not to have had an exclusively utilitarian agenda. In ‘Henry Sidgwick’s Practical Ethics: A Defence’, in which Skelton defends Sidgwick’s *Practical Ethics* from criticisms advanced by Sissela Bok and Karen Hanson among others, Skelton points out that Sidgwick does not want to

Sidgwick was attempting to bring the common sense moralists over to utilitarianism (and that the Dualism of Practical Reason occurred because Sidgwick failed to get the rational egoists to move over to utilitarianism), and I disagree that this is the complete picture of Sidgwick's aim; the utilitarianism he is offering is of a very particular nature (one that has a fundamental deontological property), and for this reason Sidgwick is not simply offering 'utilitarianism'. Whereas Skelton's work is excellent in its probing of the specifics behind Sidgwick's arguments for such important matters such as the relationship between the self-evident principles of common sense morality, and utilitarianism's own self-evident status, Skelton has a narrow view of the type of relationship that this can be, and it is an exclusively utilitarian one.

1.3.e. Robert Audi, and the Idea of a Sidgwickian/Kantian/Rossian Synthesis

The coherentist/foundationalist debate is, in many ways, an alternative (and extremely interesting) approach to exploring whether Sidgwick's work offers some kind of reconciliatory relationship between deontological and utilitarian principles. Another angle from which this might be approached is the narrower question of whether Sidgwick's intuitionism *by itself* is comprised of deontological and utilitarian moral properties, or whether it somehow demonstrates reference to both. That his intuitionism *is* of this dual nature is part of the argument I advance in this thesis. Almost all scholars (such as Antony Skelton just discussed, and David Phillips and Peter Singer discussed below) hold Sidgwick's intuitionism to be utilitarian without question²⁷¹ - and Robert Audi is also of this opinion. But the area of Sidgwick's theory that Audi considers to be unsatisfactory, and the fact that Audi proposes W.D. Ross as a solution, draws attention precisely to what I believe is the potential for Sidgwick's intuitionism to be taken to include deontological and utilitarian properties. Audi was outlined in the introduction for his own important intuitionistic method of

convince people of utilitarianism in *Practical Ethics*, and that he doesn't assume the acceptance or validity of it (p. 213)

²⁷¹ For further examples see D.D. Raphael, 'Sidgwick on Intuitionism', in, *The Monist*, Vol.58, p.3 (1974) 405-419 (Raphael also directly agrees with my interpretation of Intuitionism as representing a deontological position p.407); Ross Harrison 'Cambridge Philosophers VI: Henry Sidgwick', in, *Philosophy* Vol.71, No.277 (1996) 423-428, p.434; Alan Donagan, 'Sidgwick and Whelwellian Intuitionism: Some Enigmas', in, *Canadian Journal of Philosophy*, Vol.7, No. 3 (1977) 447-465, p.451).

combining Kantian and Rossian intuitionism with a theory of value ‘that provides unifying grounds for all of the moral principles in question’. Here, account is taken of where and how Sidgwick features in that theory.

The first thing that should be noted is that, as a follower of Ross, and of Kant in some form, Audi is *not* aiming to include elements of utilitarianism in his moral theory, at least as it is classically recognised. There is all the more reason then to think it significant that Henry Sidgwick – that renowned classical utilitarian - should be the first intuitionist to whom Audi refers at the beginning of his own account of intuitionism. Audi examines each of Sidgwick’s three kinds of intuitionism, beginning from Perceptual Intuitionism, and moving through Dogmatic Intuitionism to the final Philosophical form. This last stage, Audi points out, embodies both generalised moral principles (i.e. those found in Perceptual Intuitionism) and the more specific ones determined by the dogmatic stage, and provides a ‘method of correction of both formulations at that middle level and of moral judgements concerning particular actions’²⁷². This, Audi explains, is Sidgwick’s attempt at ‘synthesising’ the principles of common-sense morality. Audi does point out that Sidgwick recognised even his preferred type of intuitionism to be fallible, but that ultimately Sidgwick is not vastly interested in these implications. Instead, Audi states, Sidgwick was more concerned with how we can ascertain and rectify errors in moral judgements – for it is when those inconsistencies appear in our moral intuitions that ‘the need for synthesis’ becomes the most urgent²⁷³.

Clearly then, Audi’s perception of Sidgwick’s intuitionism is that within it, Sidgwick was concerned to bring to ethics a mechanism through which our ordinary and daily moral judgements can be properly assessed, and either followed or discarded accordingly. This involved Sidgwick passing those judgements through the criteria for self-evidence – yet Audi argues that Sidgwick’s ultimate use of philosophical intuitionism lies in the movement that Sidgwick makes from philosophical intuitionism to utilitarianism. So Audi concludes here that Sidgwick’s intuitionism was essentially utilitarian, as so many other writers have done.

But what is most crucial about Audi’s treatment of Sidgwick is that he recognises a *place* in Sidgwick’s intuitionism in which a deontological aspect might have been

²⁷² Audi, *The Good in the Right*, p.8

²⁷³ Ibid. p.9

helpful. Audi refers to Sidgwick's insistence that we 'remain as far as possible in the 'region of middle axioms'', before going on to lay out his own vision of combining a modified version of Ross's intuitionism with a Kantian style of moral theory, in order that we might have a better understanding of moral obligation and moral justification, and an 'enhanced ability to determine what to do, particularly where we face conflicting duties'²⁷⁴. For Audi then, Sidgwick's intuitionism could have been met by the Rossian/Kantian formulation that he himself proposes, and this would have perhaps met Sidgwick's 'call for middle axioms' more effectively²⁷⁵.

Audi's own Kantian/Rossian intuitionism is a valuable theory, but I draw attention to Audi's overlooking of Sidgwick's own use of Kant. Where Audi has utilised Sidgwick in order to show that a Kantian/Rossian intuitionism is vital to the synthesising of moral principles, he has not recognised that Sidgwick himself did this in such a way that his resulting moral system was not merely one of utilitarianism, and that it too included a crucial reference to a 'Kantian intuitionism'²⁷⁶. Had Audi identified (or focused on) this, he could possibly have traced a direct line from Sidgwick's position, through Ross, to his own Rossian/Kantian form of intuitionism. This would have been a Sidgwickian/Rossian/Kantian version of that theory. Audi's work is therefore significant in two ways: It points to the apparent need for a deontological element in Sidgwick's intuitionism, but then does not take into account the fact that Sidgwick himself supplied this. I agree with Audi on the first count, and posit that I can make a case against the second²⁷⁷.

²⁷⁴ Ibid. pp.80-82. By 'middle axioms', Sidgwick means the 'broad agreement in the details of morality which we actually find both among thoughtful persons, and among plain men who do not seriously trouble themselves about first principles' (Sidgwick, *Practical Ethics: A Collection of Essays and Addresses* (London: S. Sonnenschein: 1896), p.6.

²⁷⁵ As Audi himself states, combining the Kantian categorical imperative with 'relative closeness to moral practice of Rossian principles of duty' (emphasis added), yields the major benefits of both positions (p. 81).

²⁷⁶ Audi says that "This chapter will show how the modified Rossian intuitionism...can largely answer Sidgwick's call for middle axioms and how it can be integrated with a Kantian moral theory..." (p. 81). It is my view that Sidgwick himself integrated a Kantian moral theory.

²⁷⁷ Audi also recognises that Ross owed a lot to Sidgwick's work, despite the fundamental differences in their ethical positions (or at least, between Ross's position and the position Sidgwick is traditionally thought to have held. See *The Good in the Right*, pp.26, 29). In doing so, Audi is – perhaps inadvertently - demonstrating Sidgwick's inherent flexibility, and the fact that his applicability extends far beyond the realms of the utilitarianism with which he is so frequently and so freely associated.

1.3.f. Mariko Nakano-Okuno: ‘Sidgwick and “The So-Called ‘Discrepancies” Between Utilitarian and Kantian ethics’

The next theorist relevant to this section on possible recognitions of a Sidgwickian deontological/utilitarian reconciliation is Mariko Nakano-Okuno. Although the theories examined so far in this section are all crucial for what they do or do not say about Sidgwick’s uniting intuitionism with utilitarianism, the specific concepts of Sidgwick’s ideas on duty and on deontology itself have been largely missing. With Nakano-Okuno, and her direct comparison of Sidgwick and Kant, we move on to that most central question of where else a deontological element has been recognised in Sidgwick’s work, and whether a connection has been made between that deontology, and Sidgwick’s utilitarianism.

As suggested by the title of this 2004 paper, given at the Proceedings of the World Congress on Henry Sidgwick, Nakano-Okuno’s understanding of the relationship between deontological and utilitarian moral principles is that there is far less distance between them than is commonly imagined. She begins her argument from an observation similar to mine at the start of this thesis, that moral philosophy cannot help but refer to utilitarian or Kantian ethics, whenever it discusses the eternal questions of the meaning of right, or good, or freedom, and duty, and that the discrepancy between deontological and utilitarian principles is, to use her phrase, ‘taken for granted and hardly questioned’²⁷⁸. With these differences being, in Nakano-Okuno’s opinion, ‘overstated’, her intention here is to argue that the two schools of thought actually share several important ideas, which are ‘essential elements for both’²⁷⁹. Her paper then could be thought of being, like this thesis, an attempt to close the distance, or ‘heal the divide’, that has traditionally permeated deontological/utilitarian relations. Nakano-Okuno’s language is that of reconciliation, ‘close relationship’, and, as she states at the end, ‘constructive dialogue’ between these two giants of moral philosophy.

In terms of Sidgwick’s role here, Nakano-Okuno uses *ME* as the major text for utilitarianism, believing – as has become so much the norm – Sidgwick’s account of the doctrine to be the most sophisticated²⁸⁰. What is revealed most obviously by

²⁷⁸ ‘So-called Discrepancies’, p.260

²⁷⁹ Ibid.

²⁸⁰ The Kantian texts against which Sidgwick’s utilitarianism are compared are the *Groundwork of the Metaphysics of Morals*, the *Critique of Practical Reason* and the *Metaphysics of Morals*. Nakano-Okuno outlines the general reasons for the conflict between Kantian and utilitarian principles, listing

Nakano-Okuno's argument is just how easy it is to correlate deontological and utilitarian principles – when the latter are conceived of in *Sidgwick's* terms. Nakano-Okuno summarises Sidgwick's moral position with what she considers to be his three main objective truths – the principle of Justice (which is the Kantian maxim), the principle of Prudence, and the principle of Rational Benevolence. These give the normative ought, the temporal neutrality, and the equal treatment of all people respectively – and these, she argues, correspond to the formulae of the categorical imperative. She also points out that the traditional utilitarian concept of Happiness, or hedonism, has not yet appeared, and that when it does this happiness is defined as the Ultimate Good (which is ultimately going to coincide with Kant's view)²⁸¹. Nakano-Okuno then begins the process of examining the 'overlaps' between the fundamental principles in this position, and those in the Kantian one. She first expressly recognises that Sidgwick professed his principle of Justice to correspond directly to Kant's fundamental formula of the categorical imperative²⁸². He has some caveats about it, but ultimately, 'what the principle of Justice and the formula of universal law require is essentially the same'²⁸³. Nakano-Okuno goes as far as to state that 'the utilitarian principle itself is not a fundamental principle; rather, the validity of every utilitarian moral judgement must be tested in the light of this Fundamental Principle of Justice i.e. the Formula of Universal Law'²⁸⁴. Nakano-Okuno goes on to argue for other important parallels between Sidgwick and Kant, including a claim that Kant's Formula of an End in Itself corresponds to Sidgwick's Rational Benevolence, the fact that Sidgwick's principles do not depend on will as the determining ground but rather on prior-accepted laws, that Sidgwick's self-evident and non-tautological principles are consistent with Kant's *a priori* principles, and that Kant also included ultimate *ends that are duties* which are happiness and

the standard Kantian formulas (the categorical imperative, and the other *a priori* laws) and identifying the common grounds for the alleged conflict to lie in either a total difference in their fundamental moral principles, the grounds for the determination of moral will, the priority of right or good, attitudes to hedonism, the significance of freedom, or the fundamental principles of morals ('So-called Discrepancies', pp.264-278). It is this last point from which Nakano-Okuno's use of Sidgwick, and her reconciliation, truly begins.

²⁸¹ Ibid., pp.282, 296.

²⁸² Nakano-Okuno actually refers to the preface to the sixth edition (the PD) for support of this point, on more than one occasion. She does not, however, read the relevant passages from the PD as representative of Sidgwick's own reconciliatory idea.

²⁸³ Ibid. p.290

²⁸⁴ Ibid.

perfection, as Sidgwick also believed them to be²⁸⁵. Overall, she theorises the apparent divergence between Kant's theory and utilitarianism to be the result of the fact that Kant never combined the idea of supreme principles with that of ultimate ends, instead presenting each as independently important. Sidgwick, on the other hand, did combine them, and obtained the Greatest Happiness Principle. That is, the difference lies in the respective construction of their principles, not in the fundamental moral principles themselves.

I agree wholeheartedly with Nakano-Okuno's assessment of the correlation between Sidgwick and Kant's theories, and her work is clearly important to the cause of drawing general attention to the fact that Sidgwick was not, in many ways, simply a standard utilitarian, and that his work demonstrates significant Kantian tendencies²⁸⁶. But I do part ways with her on two counts. The first of these is her ultimate conclusion that Sidgwick's combining of the principles led to *utilitarianism*. The reason that I disagree with her on this leads to the second point of divergence, which is that Nakano-Okuno has not actually taken into account the idea of reconciliation between deontological and utilitarian principles as it appears in *Sidgwick's* own work. She has recognised that Sidgwick looked to a fundamental notion of absolute duty, and that this had its roots in the traditional Kantian notion of duty, no less. But this is as far as her discussion on the role of duty in Sidgwick's moral theory goes - and it is for this reason that I believe her to have missed the other 'close relationship' which is that between non-consequentialist duty and utilitarianism in *Rational Benevolence*. It is surprising that she has missed it, given the points that she has made. Bizarrely

²⁸⁵ Ibid., p.306. The first claim – that Rational Benevolence accords with Kant's Formula of an End in Itself - is particularly striking, mainly for its apparent simple truth. Nakano-Okuno points out that both expressly state that we should treat a person as a subject with his own ends, and that all people should be treated equally. Although I do not argue this point directly in this thesis, I agree entirely with this similarity Nakano-Okuno has drawn, and appropriate it momentarily here as further evidence of there being an essential deontological basis to Rational Benevolence.

Further to the last point, which is a particularly important one for showing the similarity between Sidgwick and Kant's positions, Nakano-Okuno argues that this sole ultimate end can only be established *after* the fundamental ethical principles in both Kant *and* Sidgwick. Not only does she then attribute Kantian concepts of duty to Sidgwick, but she also attributes a type of consequentialism to Kant. This is not the angle that I necessarily take in this thesis, but as Sidgwick says, 'no morality ever existed which did not consider ulterior consequences to some extent' (I:VIII:96), and where this should not be read as a statement that all morality is ultimately utilitarian, it can be read as further grounds on which to argue that the distance between deontological and utilitarian ethics is actually less than it might first appear.

²⁸⁶ It might be said, alternatively, that Nakano-Okuno is not so much concerned with proving Sidgwick to be a deontologist of some kind, but rather with proving Kant to be a consequentialist of some kind (see especially her discussion of ethical hedonism as an a priori principle in both Sidgwick and Kant, pp.306-308).

then, Nakano-Okuno's argument assumes the peculiar position of being at once one of the closest to my own, and yet the furthest away²⁸⁷.

1.4. Current Wider Readings and Interpretations of Sidgwick

In this last part of this narrative of how Sidgwick's work has been passed down and read through the years, the presentations of Sidgwick are somewhat broader than those discussed above. The main focus remains on where theorists have broached the possibility of reconciliation between deontological and utilitarian principles in Sidgwick's work – and on points to do with where Sidgwick is associated specifically with a deontological idea of duty, or how his utilitarianism is read and applied. But some aspects of the wider interpretations of Sidgwick are taken into account for the fact that these have had major impact on shaping the entire field of Sidgwickian studies – to which this thesis is directly relevant - as it stands today. Without consideration of these theorists' contributions, a history of Sidgwick's work would simply be incomplete.

1.4.a. Sidgwick and Peter Singer: The Point of View of the Universe

Peter Singer is possibly the most significant Sidgwick scholar of modern times. Where Derek Parfit recurses to Sidgwick's views just as often – and is equally crucial to the revival of interest in Sidgwick – Singer's many ethical writings have often been aimed at the public sphere, and as such have had wider exposure. It is also important to note that Singer addresses a wide range of popular and often inflammatory moral issues, and that he does this from a firmly utilitarian position - while utilising Sidgwick. For these reasons, Singer has been particularly instrumental in making Sidgwick relevant to contemporary ethics. The following overview of

²⁸⁷ See also Nakano-Okuno's book *Sidgwick and Contemporary Utilitarianism* (London: Palgrave Macmillan: 2011), in which she states Sidgwick's belief that he had reconciled dogmatic intuitionism with utilitarianism (pp. 25-26), although her interpretation of that reconciliation is that dogmatic intuitionism is dependent on and controlled by utilitarianism. (Nakano-Okuno also includes a brief discussion of the 'status of utilitarianism' in *ME* here, but again the main focus is on its role as the systemiser of common sense morality). I do not disagree that this is *part* of the reconciliation; it will be seen in chapter 4 that I argue that utilitarianism 'contextualises' the sort of non-consequentialist, absolute-principle foundation from which the rules of dogmatic intuitionism are built. The difference is I argue that that relationship works both ways – that utilitarianism itself depends on those non-consequentialist, absolute-principle foundations.

Singer's arguments is woefully condensed, but the point is to emphasise where Sidgwick has featured in the vast work of this controversial philosopher, and specifically to examine Singer's particular interpretation of how *duty* features in Sidgwick's work.

Singer, a pupil of R.M. Hare's, was originally associated with a form of preference utilitarianism²⁸⁸. This first became apparent in 1975 when Singer published *Animal Liberation* - his first major work in which he took the unusual step of resurrecting the Benthamite view on pleasure and pain in order to argue that the interests of sentient beings such as animals ought morally to be considered, on account that they can feel pain. His 1979 volume *Practical Ethics* saw Singer develop this argument to a position from which he argued that as preferences are tied to a being's capacity to experience pain and pleasure, they can be ranked

²⁸⁸ As outlined in the introduction, Preference utilitarianism is established by a crucial universalizability aspect, on account of the fact that one's own interests – or preferences – cannot logically be found to count for more than another's, given that one may be in the position of any of the other parties involved in a situation. Equality must naturally lead to universalizability. But Singer's interpretation of the role of universalizability is thus highly demanding – even more so, in fact, than Kant's, or Rawls'. Where Kant's categorical imperative and Rawls' original position only applied to *rational* beings (rationality and Reason being the core indicators of what we ought to do), Singer's does not. And it is this that has caused such controversy to surround Singer's work. Removing the necessity for interested parties to be rational had two major effects. The first was that it immediately brought in to the realm of moral decision making a wide range of other parties with interests, or preferences, who had not previously been thought relevant to this context, such as animals. The second effect was that it *discounted* certain other parties who conventionally would always be thought to have a stake in moral decision making: The most obvious of these parties is that of human beings who cannot reasonably be held to be *rational*, such as fetuses, infants, and the mentally disabled. To expand this aspect of Singer's moral theory, it helps to return to the core argument of *Animal Liberation* (London: Pimlico, Random House: 1975) in which Singer first defends the idea that the interests of sentient beings other than humans need to be included in the moral equation. At the root of this claim is his argument that the capacity to feel pain is often more morally significant than an animal's intelligence. Singer does not (despite the claims of many of his accusers) necessarily dictate that equal treatment should be shown to all those with interests, given that some naturally have different interests and that this warrants different treatment (for example, the interests of those with disabilities will naturally be different to those who do not have a disability (see *Practical Ethics* (Cambridge: Cambridge University Press: 1993) pp. 52-54). However, Singer does argue that his equality principle refutes moral personality, rationality, and intelligence as reasons for claims on equality (see *Practical Ethics*, pp. 17-54, for a full explanation of the equality theory), and that this being the case, equality clearly cannot be limited to humans. In terms of intelligence, whereas animals are generally of a lower intelligence than average humans, there are many humans who also have diminished intelligence, as the result of a birth defect or otherwise acquired disability. Conversely, animals such as chimpanzees have learned sign language, and thus display levels of intelligence that far surpass those of disabled humans. This equality, which naturally extends to all sentient beings in Singer's theory, brings with it the basic interest, applicable to all parties, of the desire to avoid pain. All sentient beings have the capacity to suffer, and whereas some may have the capacity to suffer more than others (chimpanzees removed from their young, for example, display far more intense distress and misery than when rabbits are removed from their young), all capacities to suffer must be taken into account. "If a being suffers", Singer states, early on in this section of his argument for equality, "there can be no moral justification for refusing to take that suffering into consideration" (*Practical Ethics*, p. 57).

according to how great that capacity is. This has led to some of Singer's most controversial claims about infanticide, abortion, animal rights, and euthanasia²⁸⁹. But the point about Singer is that whereas his views may appear extreme in some ways, on further reading, they are actually highly defensible. Singer uses an often almost impenetrable line of reasoning in all his ethical arguments, and as a result his views have largely held fast, despite the constant criticism. Throughout all of this, Singer has made reference to Henry Sidgwick.

Singer is not himself, like Sidgwick, an intuitionist²⁹⁰. However, he believes Sidgwick to be a very *particular type* of intuitionist, whose 'intuitive propositions of real clearness and certainty' are *rational* intuitions that may account for the fact that intuitions are apparently an irrefutable part of our moral experience²⁹¹. In Singer's conclusion to his critique of intuitionism in *Ethics and Intuitions* (in which he examines the most common intuitionist responses to Philippa Foot's famous 'trolley problem') he draws particular attention to the last of those clear and certain axioms of Sidgwick's, quoting directly: "The good of any one individual is of no more importance, from the point of view...of the Universe, than the good of any other". It is this phrase – 'the point of view of the Universe' – that Singer has adopted as the fundamental summary of his ethical beliefs²⁹². Singer's use of 'the point of view of the universe' is perhaps best demonstrated in his 1993 book, *How Are We To Live?*, in which Singer – writing for a public readership – propounds the importance of

²⁸⁹ It is clear why Singer's theory has attracted such debate. His views make it possible to argue that a chimpanzee could potentially have greater moral status than a human infant, and indeed Singer does argue that this could be the case. His emphasis on the capacity for suffering, and its weight in moral questions, has also led him to argue for infanticide in cases where the infant's present and future suffering far outweighs its interest in being alive, and of course for abortion when the mother is suffering in some way. His argument for abortion is doubly supported, given that the foetus does not actually possess morally relevant characteristics (*Practical Ethics*, pp. 150-152), and does not conform to definitions of 'personhood', such as rationality, self-consciousness, awareness and the capacity to feel. His critics have generally decried his 'challenge to the superiority of human life' (Singer's own term, from the preface to the second edition of *Practical Ethics*), and of course his propounding of views that so flagrantly violate certain moral standards that have for so long been accepted (such as abortion, infanticide, and even bestiality). For a full overview of the range of arguments against Singer, see the two excellent volumes Dale Jamieson (ed.) *Singer and his Critics* (Oxford: Blackwell Publishers: 1999), and J.A. Schaler (ed.) *Peter Singer Under Fire* (Chicago and La Salle, Ill: Open Court: 2009). The second volume also includes Singer's responses to each of the arguments put forward against him.

²⁹⁰ See Singer, 'Ethics and Intuitions', in, *Journal of Ethics*, Vol.9, No.3-4 (2005) 331-352

²⁹¹ Ibid., pp.350-351

²⁹² Singer no longer adheres to preference utilitarianism, having moved instead to a more classical and hedonistic utilitarian position. The 'point of view of the universe' is not only just as relevant to this position, but has at least to some extent been the inspiration for it.

adopting a broader and more inclusive attitude to our societies²⁹³. Singer introduces the phrase during the final chapter, entitled 'The Good Life', in a grand reveal of an argument that has unfolded steadily, and convincingly. Once it has been introduced, Singer uses the expression eleven times in seven pages. 'From this perspective', Singer says, 'we can see that our own sufferings and pleasures are very like the suffering and pleasures of others; and that there is no reason to give less consideration to the sufferings of others, just because they are 'other''²⁹⁴.

The point of view of the universe is also, according to Singer, a direct result of our capacity to Reason²⁹⁵. As I will be arguing here, Sidgwick establishes the point of view of the universe via Reason, and Singer does not disagree. But what is particularly interesting is that Singer does *not* associate this Reason with the production of an unconditional, non-consequentialist form of duty. In fact, Singer rejects Kant from the equation altogether, with characteristic confidence: "Let us throw out, once and for all", he declares, "Kant's idea that moral worth is to be found only when we do our duty for the sake of doing our duty"²⁹⁶. This means that Singer has explicitly divorced Kant's version of Reason and duty, from Sidgwick's version of Reason and duty. Because Singer does believe that Benevolence, the informant of

²⁹³ Singer, *How Are We To Live?* (Oxford: Oxford University Press:1993).

²⁹⁴ In *Practical Ethics*, the phrase is used only once – during Singer's closing emphasis on the importance of objectivity for acting ethically (*Practical Ethics*, p. 334). However, the words can be seen to have a steady presence in Singer's views, and in those who have engaged in discourse with Singer. During his response to Frank Jackson's article on non-cognitivism (Jackson, 'Non-cognitivism, Normativity, Belief', in, *Ratio*, Vol.12, No.4, (1999), 20-435) and Singer's apparent non-cognitivist position, Singer's first concern is to dispute that label by referring Jackson to the fact that, for him, the plausibility of Sidgwick's 'point of view of the universe' axiom is undeniable, and that this implies that there is at least one ethical judgement that *can* be known (Singer, 'A Response', in, Dale Jamieson (ed.) *Singer and His Critics*, pp. 269-335, p.270), and that he attempted to write both *How Are We To Live* and *The Expanding Circle: Ethics, Evolution, and Moral Progress* (Princeton: Princeton University Press: 2011) on its basis. Singer furthers his point against Jackson a few pages later, by arguing that the point of view of the universe is "something that as a rational being I can come to understand" ('A Response', p. 285). In a response to Michael Huemer ('Singer's Unstable Meta-ethics', in, *Singer Under Fire*, pp. 359-379) (who questions whether Singer is actually a non-cognitivist, given that he often follows Hare's arguments from moral language) on whether he is an intuitionist on the basis of his heavy appeals to Sidgwick, Singer inclines to agree that his position seems to change, but points to ultimately favouring Sidgwick's advantage in believing the element of moral universalizability to be a self-evident truth: Here, again, Singer references the point of view of the universe (Singer 'Reply to Michael Huemer', in, *Singer Under Fire*, 380-394, pp.391-392). Jan Narveson ('Singer on Moral Theory', in, *Singer under Fire*, pp. 463-487) discusses Singer's use of Sidgwick's phrase at length, pointing out that Singer has been influenced by it to the point of affording it 'the same sort of axiomatic status that Sidgwick credited it with' (p. 467). Narveson himself does not believe 'the point of view of the universe' to be an effective, or even a true, metaphor – but he links Singer with it to the extent to which he refers to it as 'the Sidgwick-Singer formula' (p. 469).

²⁹⁵ *Practical Ethics*, p.272

²⁹⁶ *Ibid.* p.215-217

the point of view of the universe, is in fact a duty – this is extremely clear in *The Expanding Circle*²⁹⁷.

So Singer takes an extremely different view of Sidgwick's use of Reason to mine, leaving out the idea of Reason establishing the sort of non-consequentialist duty for which I am arguing. But Singer does still utilise Sidgwick's version of Reason to argue for an objective ethical *right*. This is precisely what the point of view of the universe is meant to represent. The phrase is definitive of a certain aspect of Singer's moral theory, which is that his utilitarianism – and Sidgwick's - indicates an *objective* viewpoint that is the result of Sidgwick's emphasis on Reason and rationality. This objectivity is fully utilised in Singer's most recent work, *The Point of View of the Universe: Sidgwick and Contemporary Ethics* (2014) co-authored with Katarzyna de Lazari-Radek²⁹⁸. This comprehensive volume is a defence of Sidgwick's hedonistic utilitarianism, during which the authors identify challenges to this view, and compare Sidgwick's position to alternative and competitive theories²⁹⁹. Despite their firm assertion that Sidgwick is a classical utilitarianism – and their own

²⁹⁷ *The Expanding Circle*, pp.30-31, 50, See also Singer and Lazari-Radek, *Point of View of the Universe*, pp.72-74. Singer does make many other references to Sidgwick that do not specifically involve the point of view of the universe. For example, Singer argues against Narveson that all three of the classical utilitarians, including Sidgwick, naturally thought of their theories concerning pleasure and pain as extending to animals (*Singer Under Fire*, p. 460). Singer also utilises a different quote from Sidgwick at the beginning of his chapter on the biological basis of ethics in *The Expanding Circle* (p.23) but this too clearly supports Singer's strict and all-encompassing universalistic view of morality. Singer frequently draws on Sidgwick's distinction between what it may be right to privately recommend, and what would not be right to advocate openly (part of Sidgwick's esoteric argument), during the complicated discussions regarding whether utilitarianism dictates that it may be to the greater benefit if some actions are not publicly followed (see *Singer Under Fire*, p. 241, *The Expanding Circle*, p. 166, and *Singer and his Critics*, pp. 292, 301). Overall, the impression gained from even a perfunctory reading of Singer's views, and the responses to his work, is one that is imbibed with Sidgwick's presence. Of all theorists who either reference Sidgwick or have at least recognised the importance of his work, perhaps no other has attached Sidgwick to their ethical views as frequently and so concretely as Peter Singer. In a rather touching turn of phrase that Singer himself uses during a response in *Singer and his Critics*, Singer describes his source of inspiration as 'my philosophical hero, the far-seeing and by no means conservative, Henry Sidgwick' (p. 314)

²⁹⁸ Singer and Lazari-Radek also give a strong defence of Sidgwick's use of Reason to produce ethical objectivity and impartiality (embodied by Universal Benevolence) in 'The Objectivity of Ethics and the Unity of Practical Reason', in, *Ethics* Vol. 123, No.1 (2012) pp.9-31

²⁹⁹ Singer and Lazari-Radek accept absolutely that Sidgwick intended to give a normative theory in *ME*. As I have outlined in the introduction, I do not entirely share this view, on the basis that Sidgwick specifically states in the first and second prefaces to *ME*, and in the introduction, that his aim is not to discuss what people *ought* to do, but rather what they *do* do. I do agree with Singer and Lazari-Radek however, that Sidgwick's treatment of the notion of 'ought', and its inherent relationship to Reason, does eventually yield a normative element, but that this – and his moral theory overall - arose inevitably from Sidgwick's investigation, at odds with those originally stated intentions. I return to this point in chapter 4.

Singer and Lazari-Radek argue their case for this hedonistic understanding of utilitarianism on the basis of Sidgwick's argument that the sole ultimate good is 'desirable consciousness', and that this is pleasure/happiness (*Point of View of the Universe*, pp. 207-212).

self-confessed intention to defend utilitarianism themselves - the authors want to emphasise that ethical judgements are objective truths we can know by Reason³⁰⁰. Utilising Sidgwick's account of moral motivation, they argue that this begins with the rational judgement that an act is right, and that this has concomitant impulses for doing that action³⁰¹. 'This account of moral motivation', they say, 'preserves the objectivity of moral judgements and their ability to motivate'³⁰².

The authors thus intend to bring a Sidgwickian account of the role of Reason to ethical theory and practice, and in doing so they have retained the Reason that I believe leads Sidgwick to a deontological version of duty very similar to Kant's. But they specifically eliminate Kant as an influence. Although this perhaps presents a significant alternative reading of the type of duty in Sidgwick's work, I do not see how Singer and Lazari-Radek can rely so far on Sidgwick's definition of rightness for their theory of moral obligation when this is so bound up with a version of Reason and rationality so similar to Kant's, and then entirely reject the Kantian notion of duty from the theory of moral obligation that Sidgwick eventually builds from that model of

³⁰⁰ Singer is slightly reticent about defending moral objectivity in *How Are we to Live?*, stating "I am not defending the objectivity of ethics in the traditional sense. Ethical truths are not written into the fabric of the universe (p. 275). In *Point of View of the Universe*, Singer and Lazari-Radek have made a full move to defending objectivity. Sidgwick's point of view of the universe, as is evident from the title, remains the foundation of that objectivity.

³⁰¹ Ibid, p.64. Other inclinations such as sympathy and moral dissonance are also important motivating factors, but Reason is the master for Singer and Lazari-Radek's Sidgwick (*Point of View of the Universe*, pp. 62-66).

They also discuss whether Sidgwick's foundationalism can act as an alternative to the method of reflective equilibrium favoured by coherentist theories (p.94). Their conclusion here is that Sidgwick's 'point of view of the universe', which is rooted in his emphasis on rationality, gives to morality an objectivity that reflective equilibrium, or other coherentist theories, cannot (from this basis, they are prepared to argue definitely, where Sidgwick was only tentative, that 'the point of view of the universe' is the perspective of a rational being, where egoism is not (p. 378)). I return to Singer's objections to Sidgwick's 'reflective equilibrium' shortly. In terms of the foundationalist/coherentist debate explored above, Singer and Lazari-Radek refer to Robert Audi's distinction between strong and modest forms of foundationalism (p. 108). As modest foundationalism allows a role to coherence, this may be the most appropriate way of understanding Sidgwick to be at once a foundationalist and a coherentist. This does not quite accord with Audi's distinction however, which holds strong foundationalism, or 'classical ethical foundationalism' to refer to those theorists who were seeking to establish self-evident principles that were 'directly justified and un-revisable', whereas the modest ethical foundationalist in Audi's distinction only holds some ethical beliefs to have prima facie value/justification. It is immediately apparent from this more detailed explanation of the two positions that Sidgwick would belong in the first camp, and Ross, for example, in the second. The extent to which Sidgwick and Ross can both be called foundationalists under Audi's distinction, and how far the respective work of those two philosophers can be made 'coherent', in the non-ethics sense of the word, would be the subject of further work.

³⁰² They also argue that Sidgwick's search for rational justification, and the resulting hedonistic utilitarianism, supersedes Ross's prima facie duties as a means of deciding between moral duties.

Reason³⁰³. It might be argued that the point of view of the universe is simply a wholly utilitarian concept, with utilitarian aims and values, and I admit that this is a legitimate interpretation. But this still does not necessarily preclude the idea of that view being established as a categorical and non-consequentialist type of duty. On such a reading it would *still* be a utilitarian theory, and it would still incorporate the hedonistic utilitarianism of Singer and Lazari-Radek, but it would have the crucial quality of strong moral bindingness. This seems to be how Singer and Lazari-Radek want to present the point of view of the universe (given that they use the phrase to represent a normative ethical theory) – and it is certainly how Sidgwick himself presented it. But when they do present Sidgwick’s form of Rational Benevolence, it is from an intuitional basis, not a deontological one. I do not dispute that intuitionism is of course a key epistemological method in the establishment of this principle, but I approach this point from a different angle by arguing that Sidgwick’s intuitionism and Sidgwick’s Intuitionism are very similar when it comes to Sidgwick’s use of Kant, and that therefore – even where epistemological intuitionism *is* key - there is *still* a deontological basis to that Rational Benevolence³⁰⁴.

³⁰³ In terms of their interpretation of the relationship between Kant and Sidgwick, they recognise that Kant’s maxim was important to Sidgwick’s principle of Justice – but argue that this maxim is the not the sole criterion of moral rightness. On this, I agree with them – rightness, embodied by Rational Benevolence, also requires Prudence in order to generate Universal Benevolence (and thus the point of view of the universe). But I do not believe this gives grounds on which to overlook it, and I do not believe Sidgwick does overlook it. (They also concede that one could draw a parallel between Sidgwick and Kant’s statement that ‘duty is the necessity of an action from respect for the law’ (*Groundwork for the Metaphysics of Morals*, 16; *Critique of Practical Reason*, 1:6:6)

³⁰⁴ Singer and Lazari-Radek do go some way towards acknowledging that the non-consequentialist aspect of Intuitionism may appear again in Sidgwick’s epistemic intuitionism (*Point of View of the Universe*, pp. 69-70). They first draw attention to the initial confusion that seems to arise over the use of ‘intuitionism’ to refer to the content of our moral judgements, and the use of the term to describe the process through which we reach moral judgements. But there is, they state, a philosophical connection between intuitionism in the first, wider, epistemic sense (i.e. that which I distinguish with a small ‘i’) and intuitionism in the narrower, ‘common sense morality’ sense (that which I distinguish with a capital ‘I’) ‘with its demand for obedience to rules irrespective of consequences’ – which is that if a theorist is going to maintain that certain truths are obtained just by looking at them, then this must exclude reference to consequences, or this ceases to be real intuitionism. The authors quote ME III:I:200 in support of this, which I also use in the next chapter as evidence for my argument that Sidgwick’s Intuitionism (in the narrower sense) represents the traditional non-consequentialist form of deontology, *and* for support of a point that Singer and Lazari-Radek have themselves just made, but not (in my view) expanded. This is my argument that Sidgwick himself, as an intuitionist (in the wider, epistemic sense, with a small ‘i’) *also* adheres to some extent to a position in which consequences are not relevant to the knowledge of certain moral truths (i.e. to Intuitionism in the narrower sense) and that therefore his moral theory is *built* from this position. Singer and Lazari-Radek have recognised the same connection between the two forms of intuitionism that I do, but they have not – explicitly at least - taken this final step of recognising that if Sidgwick is an intuitionist in the wider, epistemic sense, then on his own definition he must also be an Intuitionist in the narrower sense.

The main point from which I diverge with Singer and Lazari Radek then is this issue of the role and *nature* of the duty in Sidgwick's moral theory³⁰⁵. 'Duty' for them clearly refers to the wider sense of the word, in which it applies just as much to utilitarianism (or any other moral position) as it does to deontological positions; I maintain that 'duty' as it appears in Sidgwick is the specifically non-consequentialist form of duty, found in the deontological tradition.

On the issue of a possible reconciliation between deontological and utilitarian principles in Sidgwick's work however, Singer and Lazari-Radek's position is possibly the closest to mine of all theorists in this chapter. In a chapter entitled 'The Profoundest Problem of Ethics', they specifically recognise that Sidgwick thought the antithesis between utilitarianism and common sense morality can be overcome, and that this happens because whereas non-self-evident common sense morality must appeal to the principle of utility, utilitarianism itself 'must ultimately rest on an appeal to principles or axioms that we intuitively grasp as self-evident...Intuitionism is not really incompatible with utilitarianism; rather, intuitionism leads to the self-evident moral principles that form a rational basis for utilitarianism'³⁰⁶. It is the allusion to there being a *reciprocal* relationship between Intuitionism and Utilitarianism in Sidgwick's thought that creates a strong parallel here between their view and mine. They draw attention to the fact that utilitarianism itself looks to something else for its support, and in doing so do not simply discard common sense morality, or the intuitionism therein, in favour of a utilitarianism that can simply leave all other modes of moral reasoning behind³⁰⁷. However, this is an instance in which I believe a

³⁰⁵ Singer and Lazari-Radek do however agree that Sidgwick's version of Intuitionism is a form of deontology understood in the traditional form which I take it to mean (*Point of View of the Universe*, p. 24).

³⁰⁶ Ibid, p.149. The 'profoundest problem of ethics' to which they refer is of course the conflict between rational benevolence and rational egoism, Sidgwick's Dualism of Practical Reason. Singer and Lazari-Radek make the unusual step of claiming that they can solve this via an evolutionary argument. (David Phillips also argues that Sidgwick was more successful in overcoming this problem than Sidgwick himself realised. See *Sidgwickian Ethics*, pp.134-151). They appeal to the fact that there are better grounds for supporting an intuition for which there is no evolutionary explanation than there are for supporting an intuition that *can* be assumed to have an evolutionary origin, and posit that this throws egoism into serious doubt (p. 195). This being the case, Singer and Lazari-Radek then tentatively conclude that all reasons for action are impartial, which thus supports the 'point of view of the universe' and rather neatly resolves the dualism in favour of benevolence rather than egoism.

³⁰⁷ In the reflective equilibrium paper, after having shown that Sidgwick does not use that method, Singer asks what then is the significance of the 'coincidence of utilitarianism and common sense morality' ('Sidgwick and Reflective Equilibrium', p. 507) It means, for Singer, that utilitarianism is doubly supported - by its initial plausibility and the absence of any inconsistency or indeterminacy. It also receives the additional support of the method of intuitionism. "The result of unfolding the method

demarcation between Intuitionism and intuitionism is essential, both for clarity, and for substance. Singer and Lazari-Radek have stated that intuitionism leads to the self-evident principles that form the rational basis for utilitarianism, but this is not the same thing as saying that *Intuitionism* - understood as representing the deontological position- leads to those principles. If we *were* to recognise that it is Intuitionism in the methodical rather than just the epistemic sense that Sidgwick means here, then the reconciliation between the deontological aspect of common sense morality and utilitarianism is made much more robust. Again, this depends on understanding Sidgwick's method of Intuitionism as being in part the same as his epistemic-device-intuitionism.

On a final note regarding the idea of reconciliation in Singer's work, there is the issue of reflective equilibrium – a cause with which Singer became notably involved in the 1970s when Rawls argued that Sidgwick was using this method, and Singer directly challenged him. As I noted above, reflective equilibrium is closely connected to coherentism, and as such this method of obtaining moral judgements could also possibly be put forwards as a means of reconciling deontological and utilitarian principles and impulses. But Singer disputes that Sidgwick was using reflective equilibrium as a method of testing the validity of moral judgements, and I agree with Singer's on this, for two reasons³⁰⁸. The first is weaker; I agree with Singer that

happened to be that the method of intuitionism and the method of utilitarianism are shown to be reconcilable, because the judgements of common sense morality turn out to require the utilitarian principle to fall back on as an underlying self-evident first principle" (p. 505) Singer also refers to the preface to the sixth edition (the PD) to support some of his points made in this paper, but he does not read the passages in the way that I do.

³⁰⁸ As previously mentioned, Rawls drew on Sidgwick's own careful balance between common sense morality and the fact that that common sense morality is not absolute, fixed, or capable of real proof, to support his own argument that when there is no inherently plausible theory that matches our immediate intuitional moral judgements, then we need to modify the theory or the judgements until there is an equilibrium between them (see also Singer, 'Ethics and Intuitions', p. 344). Using Schneewind's account of reflective equilibrium in Sidgwick ('First Principles and Common Sense Morality in Sidgwick's Ethics') Singer refers to the passage in *ME* in which Sidgwick states that "perhaps we may say that what is needed is a line of argument which on the one hand allows the validity, to a certain extent, of the maxims already accepted, and on the other hand shows them not to be absolutely valid, but needing to be controlled and completed by some more comprehensive principle" (IV:II:420). The first principles are those of common sense morality, and the comprehensive principle is that of utility. And it is, as even Singer admits, reasonable to assume that this is a form of reflective equilibrium. But Singer does not himself share the view that this was Sidgwick's method for testing moral theories. Singer, who believes Sidgwick to be 'among the clearest and most careful thinkers in his area' examines Sidgwick's methods of argument, in order to show that whereas Sidgwick appears to be arguing *for* the utilitarian principle, he is actually using utilitarianism as an example of the sort of *ad hominem* argument that might be adopted if one were trying to prove the superiority of one principle over another. Ultimately, Singer rejects the idea that Sidgwick was using reflective equilibrium, because Singer rejects the idea that intuitions can be used for moral justification

Sidgwick's focus was on the establishment of truly valid self-evident first principles, but this means I would have to agree that Sidgwick was a foundationalist, and as discussed above I do not believe entirely that Sidgwick was a foundationalist (as Singer does). On the second count, my agreement with Singer is firmer. This is that reflective equilibrium, on Singer's view, is actually subjectivist. Singer does not believe Rawls has left room for validity that is independent of achieving reflective equilibrium, and therefore the validity of a moral theory will 'vary according to whose

(see his emphasis on rationality as outlined above). He is thus drawing attention to the fact that Sidgwick's approach was far more scientific than that, and that he therefore produced a far more objective system of justification. (See Eivind Kirkeby's 'Against Moral Intuitions', (https://www.duo.uio.no/bitstream/handle/10852/24997/Against_Moral_Intuitions.pdf?sequence=1) (2009) accessed 8th July 2016) in which he argues that there is an apparent inconsistency between Singer's rejection of intuitionism, and his claim that Sidgwick does *not* use reflective equilibrium, precisely because he is an intuitionist (p.18). Kirkeby points out that Singer attaches this intuitionism to Sidgwick's search for self-evidence, and the sort of 'first principles' which Sidgwick did not believe to be capable of being proven (given that this would deflect self-evidence onto that which proves the principle) p.19. But Kirkeby also states that once Sidgwick has argued for the self-evident axioms (Justice, Prudence, and Universal Benevolence), he then argues that these need to be tested against our commonly held principles of common-sense morality (p.23). Singer still denies that this a form of reflective equilibrium, because Singer's issue with Rawls' version of this method is that Rawls assumes its results to be absolutely valid, whereas Singer does not believe Sidgwick's method to hold that the results are valid in this way. Kirkeby's description of the Sidgwick/Singer/Rawls/reflective equilibrium debate seems to me to perfectly summarise the reasons why it is so difficult to conclude on whether or not Sidgwick was using a form of that method. His self-evident principles are at the very core of his moral theory – but so is a vital coherence between those self-evident principles (or the Universal Benevolence they combine to produce) and the moral rules that we already hold (that make up common sense morality). Self-evidence seems to deny reflective equilibrium – negotiating a relationship between self-evidence and commonly held moral rules seems to absolutely confirm it. In terms of my own interpretation of Sidgwick on this point, I have said in the main text that on the basis of Sidgwick's self-evidence and his specific aim to avoid subjectivism I agree with Singer that reflective equilibrium is not Sidgwick's primary method. However, some aspects of my argument for Rational Benevolence do also leave Sidgwick on the uncertain ground between the two views. My argument for the positive proof between utilitarianism and the deontological basis of the self-evidence principles (see chapter 3) refers to that passage at IV:II:420 that Singer admits (and that I agree) appears to point towards reflective equilibrium. There is also the sister passage at III:XIII:379, to which I have already referred in the thesis introduction, in which Sidgwick talks about the truth lying between the two conclusions that there *are* certain absolute principles, but they require systemisation by another, more comprehensive, principle. I refer to these passages collectively as 'the connection passages', and they form an integral part of my argument for the reconciliation. This would seem to require that I state Sidgwick to be using reflective equilibrium. On the other hand, Rational Benevolence itself is not – on my argument – reflective equilibrium because its construction is far more solid than that; it takes as a foundation the non-consequentialist and self-evident principles of Justice and Prudence, and through them establishes Universal Benevolence, which then fully contextualises those principles within one holistic moral theory. It also seems as though 'rationally benevolent' actions (see thesis conclusion) could not be established on a basis that is still subjectivist. Furthermore, if we limit the concept of a reconciliation in Sidgwick's work to that between common sense morality and utilitarianism in this practical 'reflective equilibrium' sense, then there would be no need for the synthesis in Rational Benevolence anyway. On this basis then, I am generally persuaded to argue that Sidgwick's synthesis at least does *not* rely on reflective equilibrium). See also Steven Sverdlik, 'Sidgwick's Methodology', *Journal of the History of Philosophy* Vol. 23, No.4 (1985), 537-553, and Marcus Singer's 'The Methods of Justice: Reflections on Rawls', *Journal of Value Enquiry*, Vol. 10, No.4 (1976), 286-316).

considered moral judgements the theory is tested against'³⁰⁹. This was the problem that Sidgwick perceived in Kant (Rawls is a neo-Kantian), and it is fair to assume that Sidgwick would have wanted to avoid the same problem – hence his quest to establish truly self-evident principles. I am, therefore, following Singer's lead, and (for the most part) rejecting reflective equilibrium as representative of Sidgwick's theory – specifically as being representative of the method through which he reaches the reconciliation between deontological and utilitarian principles in Rational Benevolence³¹⁰.

In the wake of this extensive and comprehensive reading of Sidgwick presented by Singer and Singer and Lazari-Radek, there is one last point to be made. This is that as vital as all this work has been for rehabilitating Sidgwick into the modern day, and to drawing out Sidgwick's insightful and often progressive views for the purposes of modern ethics when for so long *ME* has been regarded as almost impenetrable, Sidgwick's utilitarian status is *still* - even at this late stage in his history - being reinforced, perhaps more than ever. This is particularly significant for my argument here, given that the point of view of the universe has been the foundation for so much of Singer's work and that I do not agree this position needs to be a completely utilitarian concept³¹¹. It is my view that even Singer's extensive and sophisticated reading of Sidgwick has, in its far-reaching scope, overlooked a smaller but critical piece of evidence that Sidgwick's work is *not* just classical utilitarianism.

³⁰⁹ 'Sidgwick and Reflective Equilibrium', p.494.

³¹⁰ Brad Hooker's more recent defence of reflective equilibrium as a process for obtaining moral judgements offers a more substantial and developed version than Rawls', and thus might avoid Singer's objection to the method (see Hooker's 'Rule-Consequentialism, Incoherence, and Fairness', 'Ross-Style Pluralism vs Rule-Consequentialism', and 'Reflective Equilibrium and Rule Consequentialism' in, Hooker, Mason and Miller (eds.) *Morality, Rules, and Consequences*, pp.222-238). This is a type of rule-consequentialism that does not presume a consequentialist framework, and Rational Benevolence – and the way in which utilitarian principles and the basis of common-sense morality inform each other therein - also does not presume a consequentialist framework. Further exploration of the relationship between Sidgwick and Hooker's version of reflective equilibrium in the future would therefore likely be beneficial to widening an understanding of the relationship Sidgwick saw between common sense morality and utilitarianism, but this relationship is not my specific focus in this thesis. Here I am involved with establishing Rational Benevolence as a moral theory which unifies in a meta-ethical way deontological and utilitarian principles, and for this context I generally follow Singer's subjectivist objection (despite the fact that some of my argument also appears to point towards reflective equilibrium).

³¹¹ Singer also believes that *ME* has been misinterpreted, because it has always been read as a defence of utilitarianism ('Sidgwick and Reflective Equilibrium', pp. 502-503, 505)

1.4.b Sidgwick and David Phillips: Intuitionistic Utilitarianism, and The Unfairness Objection

Phillips' relatively short but extremely important book *Sidgwickian Ethics* is testament to the now steadily growing resurgence of interest in Sidgwick's work³¹².

Acknowledging ME's 'mixed reputation', Phillips claims to offer an approach to Sidgwick that 'accentuates the positive and minimises, if it does not eliminate, the negative'³¹³. In just four concise chapters, Phillips covers all the major areas of discussion; Sidgwick's meta-ethics, his epistemology, and, most importantly for the purposes of this chapter, the two main conflicts of *ME*; as Phillips calls them, 'Utilitarianism Versus Dogmatic Intuitionism', and 'Utilitarianism Versus Egoism'³¹⁴. There is unfortunately not room here to do full justice to all of Phillips' detailed and innovative explorations of Sidgwick's meta-ethics and epistemology, and for its relevance to my argument for Sidgwick's reconciliation between deontological and utilitarian principles, Phillips' discussion of the conflict between utilitarianism and Intuitionism is obviously of particular interest. However, a few points from Phillips' earlier arguments about Sidgwick's method of epistemology are crucial to an understanding of that discussion, and also serve to give some further ground to my own arguments – especially for my claim that there is a distinct Kantian/deontological influence in Sidgwick's work. These points are discussed first. This is followed by a discussion of Phillips' 'unfairness objection' brought against Sidgwick, which also features an important contribution on this subject from Thomas Hurka.

³¹² *Sidgwickian Ethics* (Oxford: Oxford University Press: 2011). Phillips himself states that his approach to interpreting and evaluating the central arguments of *ME* is probably closest in spirit to C.D. Broad, but points out that Broad's 80-year-old treatment of *ME* 'cannot be informed by the lifetime of work in ethical theory and Sidgwick interpretation since its publication' (*Sidgwickian Ethics*, p. 8).

³¹³ *Ibid.* p.4

³¹⁴ Phillips believes Sidgwick to deserve far more attention than it has thus far received, and to be 'rich in implications for contemporary controversies' (*Ibid.* p. 7). He claims that his work is not idiosyncratic, and he is right: Although Singer and Lazari-Radek have since followed with *Point of View of the Universe*, which also seeks to apply Sidgwick to contemporary ethics and arguably does this more obviously than Phillips, Phillips' observations and arguments about Sidgwick's work are, in most places, highly distinctive. He is also one of the only theorists to recognise the importance of Sidgwick's supplementary comments made regarding the two conflicts in the personal document. He references the document early in his introduction, and specifically the sections on which this thesis relies for some of its most solid proof. These passages are used differently by Phillips however. Phillips states that Sidgwick actually believed he had overcome the conflict between utilitarianism and intuitionism, but argues that Sidgwick was less successful than Sidgwick thought. Contrarily, especially to Sidgwick's own opinion, Phillips believes that Sidgwick *did* successfully overcome the conflict between utilitarianism and egoism, the infamous Dualism of Practical Reason.

1.4.b.i 'Intuitionistic Utilitarianism'

Firstly, Phillips places great emphasis on Sidgwick's nonnaturalism, which he believes to indicate Sidgwick's status both as a foundationalist and his commitment to intuitionism³¹⁵. This intuitionism is, for Phillips, at the centre of Sidgwick's meta-ethics and his epistemology – and will result in what Phillips refers to as Sidgwick's 'intuitionistic utilitarianism'. But what is most crucial about Phillips' view of Sidgwick's intuitionism is that he specifically recognises the connection between Kant, and Sidgwick's intuitionist theory of moral obligation. This creates an important parallel between Phillips' interpretation of that relationship, and the one for which I am arguing. Phillips acknowledges that Sidgwick allows a significant role to the Kantian view of moral motivation, and argues that for Sidgwick, the Kantian view is an important and active corollary of the 'fundamental distinction' (that 'ought' or 'right' means something fundamentally different to all other concepts), rather merely being a by-product of an adherence to Kant on Sidgwick's part³¹⁶. This corresponds very

³¹⁵ Ibid. pp.12-22. See also pp. 6, 11, 52-55, 57. Phillips considers this nonnaturalism to be supported by 'the fundamental distinction thesis' (pp.12, 27, 33), which refers to Sidgwick's statement that 'the fundamental notion represented by the word 'ought', or 'right', is...essentially different from all notions representing facts of physical or psychical experience' (see *ME*: I:III:25). From this, Phillips argues, Sidgwick infers a definite 'autonomy of ethics' (p. 55 – Phillips refers to Nicholas Sturgeon for this phrase, see Sturgeon, 'Ethical Intuitionism and Ethical Naturalism', in, Phillip Stratton-Lake (ed.) *Ethical Intuitionism* (Oxford: Clarendon Press: 2002) pp.184-211) which gives intuitionism this place of absolute importance in Sidgwick's epistemology as conceived by Phillips (see also pp. 52-55, 57, 60). So crucial is the intuitionism which arises from Sidgwick's nonnaturalism on Phillips' view in fact, that it is this that becomes 'central to Sidgwick's rejection of the received distinction between intuitionism and utilitarianism' (p. 22). This also, naturally, leads to Phillips' foundationalist reading of Sidgwick (pp. 52-53), although Phillips concludes, borrowing a phrase from Audi, that this is a moderate foundationalism.

I agree with Phillips that Sidgwick's work has a nonnaturalist basis, particularly as it is this that lends Sidgwick's work the heavy emphasis on self-evidence and rationality, both of which are crucial to the moral theory for which I am arguing. I also agree that intuitionism is indeed crucial to Sidgwick's reconciling of that divide, although I emphasise the deontological nature of this intuitionism (in chapter 2), and of course argue, contra Phillips, that that reconciliation is successful.

³¹⁶ Ibid. pp.24-27, 32-33. 'The right answer, I think' Phillips says, 'is that Sidgwick's central commitment is to the fundamental distinction thesis about rationality' (p. 33) Also, Phillips says, 'Sidgwick is independently committed to the fundamental distinction thesis: 'it is more to him than just a premise in an optional argument for the Kantian view of moral motivation' (p. 24). See main text below. Robert Shaver disagrees with Phillips, and with me, that Sidgwick is committed to this fundamental concept of 'right', or 'ought'. His interpretation of Sidgwick's discussion of rightness and Reason is that this is only an optional view of Reason as a motivator. Phillips points out that Shaver's view here is then problematic for the link between nonnaturalism and intuitionism, which Shaver does deny. See Shaver, 'Sidgwick's Minimal Meta-ethics', in, *Utilitas*, Vol.12, No.3 (2000) 261-277, and 'Sidgwick on Moral Motivation', in, *Philosophers Imprint* Vol.6, No.1 (2006).

.....but he does not expand upon it specifically, or take the view – as this thesis does – that Sidgwick can in fact be called a deontologist to some extent, and that this deontology fundamentally informs Sidgwick's own moral theory. It is the task of chapter 3 of this thesis to develop this argument.

closely with my point that Sidgwick's own emphasis on rationality led him to grant a role to Reason very similar to that of Kant's, through which rightness cannot be defined by reference to consequences - and furthers the grounds for my claim that there is a traditionally deontological element in Sidgwick's work.

Phillips offers further defence for Sidgwick's sympathy with the deontological position, and for the importance of his fundamental notion of 'right'; he concludes that Sidgwick is meta-ethically committed to Kant's view of Reason as a motivator, and that, as compared to Moore, Sidgwick's fundamental notion is that of 'right', not 'good'³¹⁷. Perhaps most strikingly, Phillips is explicit that some of Sidgwick's meta-ethics *are* in fact Kantian, and that this appears in Sidgwick's argument that moral judgements 'involve commitment to the categorical imperative', as well as in his belief that rational beings are motivated by rational concepts. Phillips thus presents a strong and confident case for the Kantian influence in Sidgwick's work which I hope to continue, and expand upon here.

It is also the case that Phillips directly links this Kantian influence to Sidgwick's nonnaturalist intuitionism, as I do. On this issue of Sidgwick's intuitionism, I am mainly in complete agreement with Phillips. I agree with Phillips that Sidgwick's intuitionism is foundationalist, and I certainly agree, as I have stated, that this is a direct result of Sidgwick's commitment to Phillips' 'fundamental distinction' thesis. I also agree that it is these premises that lead Sidgwick to his reliance on self-evidence³¹⁸. Furthermore, Phillips also directly argues that it is Sidgwick's

³¹⁷ *Sidgwickian Ethics*, pp.32-38. Phillips summarises here, 'for Sidgwick, the fundamental notion is the notion of obligation' (p.36) and 'Nonnaturalist meta-ethics is better to start with Sidgwick than with Moore' (p. 38).

Phillips is also in agreement with my view that for Sidgwick, the self-evident notion of Justice is, to borrow Phillips' terminology, 'the fundamental ethical notion, a conceptual truth' (p. 42).

³¹⁸ *Ibid.* p.60. Phillips also presents a very similar but admittedly more specifically worded version of what I earlier suggested is a foundationalist/coherentist hybrid in Sidgwick's work. Where I used the term foundationalist/coherentist hybrid to refer to the wider context of Sidgwick's fusion of utilitarianism with dogmatic intuitionism, Phillips' hybrid refers to the narrower idea of whether Sidgwick's epistemological method *itself* is a hybrid of two methods: Phillips isolates the meeting between Sidgwick's intuitionistic self-evidence and his introduction of a criteria to test for that self-evidence as causing a hybrid state of affairs in Sidgwick's epistemology (p. 60). His view of this is that "Sidgwick's actual view then shares with foundationalism the idea that self-evidence has an essential epistemic role. But it also features two characteristically coherentist ideas: the idea that apparently self-evident propositions are not immune to doubt or questioning, and the idea that my justification for believing apparently self-evident claims comes in part from their coherence with other apparently self-evident claims I believe, and in part from the consistency of my beliefs about these claims with other people's beliefs about them' (pp. 61-62). (Phillips also goes on to argue that on one hand Sidgwick uses language of self-evident first principles, and on the other that he uses only bipartite and criterial arguments for the proof of the utilitarian principle, which seem to require a commitment to foundationalism and a coherentist element respectively). In terms of a possible reconciliation between

intuitionism that leads him to his ‘distinctive philosophical position’, which is an overcoming of the ‘received dichotomy between intuitive non-utilitarianism, and inductive utilitarianism’³¹⁹. For support of this Phillips quotes passages from the personal document on which I rely in later chapters as proof that Sidgwick reconciled the divide between utilitarianism and intuitionism³²⁰. As has been seen, Phillips is not the first theorist to have recognised Sidgwick’s tackling of that dichotomy - but he is the first to frame it in these specific terms.

This point, then, corresponds almost directly with my central claim in this thesis. But there are two lines of argument that I deploy in response to the challenge that this might pose. The first is that Phillips, like other theorists before him, has not taken the idea of Sidgwick’s overcoming that dichotomy far enough. This, to me, is evident in Phillips’ omission of the crucial last part of that passage from the PD. For where Sidgwick has said, as Phillips cites, that ‘the Utilitarianism of Mill and Bentham seemed to me to want a basis; that basis could only be supplied by a fundamental intuition’, he continues immediately with, ‘on the other hand, the best examination I could make of the Morality of Common Sense showed me no clear and self-evident principles except such as were perfectly consistent with Utilitarianism’ (xxi). What Sidgwick is saying here is not just that utilitarianism was in want of a prior basis of

utilitarian and deontological principles, both of those coherentist ideas are important, but the first (that self-evident propositions are not immune to doubt or questioning) is where a meta-ethical reconciliation would lie (as opposed to the more practical sort of ‘reconciliation’ between beliefs offered by a method such as reflective equilibrium. I therefore believe that if there are two such coherentist elements in Sidgwick’s moral theory, as Phillips suggests, that that first one – the meta-ethical need for coherence – is the more important, at least for the theory I aim to develop here. I am not, however, necessarily denying that Sidgwick does use that sort of wide reflective equilibrium represented by Phillips’ second coherentist element – the idea that justification for self-evident principles comes in part from coherence with other peoples’ beliefs about them. Sidgwick is often concerned with the idea of there being consensus on principles (I:I:6, I:III:27, I:VIII:100, III:I:210, III:215, and of course there is Sidgwick’s fourth criterion for testing self-evidence – universal or general consent, III:XI:341, which Phillips identifies as one of the coherentist elements). However, Sidgwick also believes that the ‘function’ of the philosopher is to ‘do more than just define and formulate the common moral opinions of mankind. His function is to tell men what they ought to think, rather than what they do think; he is expected to transcend Common Sense in his premises, and it allowed a certain divergence from Common Sense in his conclusions’ (III:XIII:373). The ‘conclusions’ that Sidgwick is referring to here, are his self-evident principles, which this passage suggests he holds to be unassailable. Phillips’ work, then, here leads to further debate over whether or not Sidgwick relies on a form of reflective equilibrium, and again the answer is unclear.

³¹⁹ *Sidgwickian Ethics*, p.57.

³²⁰ For the passages in question see the personal document, xxii-xxiii, where Sidgwick states that ‘I was then a utilitarian again, but on an Intuitional basis...I could find no real opposition between Intuitionism and Utilitarianism. The Utilitarianism of Mill and Bentham seemed to me to want a basis: that basis could only be supplied by a fundamental intuition’. These passages have a role in chapter 2 regarding Sidgwick’s ‘intuitionism/Intuitionism’, and a crucial one in chapter 4 as evidence of his reconciliation.

justification, but that the utilitarian principle, once revealed, then confirmed the self-evident principles (Justice, Prudence, and Universal Benevolence) that Sidgwick had retained from common sense morality thus establishing the two sources of morality as existing in what I will argue to be a mutually reciprocal relationship. That argument I develop in chapter 4; here my point is that Phillips has not seen the full reconciliatory value of that passage.

Phillips does, however, have another argument for Sidgwick's reconciliation, and again it is similar to mine. Phillips's view is that Sidgwick believed himself to have reconciled utilitarianism with intuitionism on the basis that there is no epistemic difference between them – as Phillips says, paraphrasing Sidgwick, 'all sensible moralists will be intuitionists'³²¹. This is, again, very similar to my argument that Intuitionism (understood as the traditional deontological position) is actually a crucial part of Sidgwick's moral epistemology. But there is a crucial difference, and it is represented in my capitalising of the word, where Phillips makes no such distinction – despite himself having recognised that there needs to be a delineation between the 'moral-theoretic' and the epistemic versions of intuitionism³²². That difference is that where Phillips believes Sidgwick to have (thought himself to have) reconciled utilitarianism and non utilitarianism on the basis of their shared appeal to intuition (with which I do not disagree), I believe Sidgwick to have reconciled them on the basis that Intuitionism (with a capital 'I', as representative of the narrower, 'moral-theoretic' sense) is also a feature of Sidgwick's intuitionism (with a lower case 'i', representative of the epistemological device sense). As Phillips says, when it comes to the moral-theoretic conflict between utilitarianism and intuitionism (which on my interpretation should be capitalised in this context), he believes Sidgwick to have resolved it with a victory for utilitarianism. That is, Sidgwick's moral theory, and his intuitionism, emerge as wholly utilitarian on Phillips' view – whereas on mine, Sidgwick's moral theory emerges as partly utilitarian, and partly *deontological*. In this way, this kind of reconciliation would actually resolve or at least address both the epistemic *and* the moral-theoretic conflict between utilitarian and non-utilitarian positions.

³²¹ *Sidgwickian Ethics*, p.53

³²² *Ibid.* p.36

I do not, therefore, agree with Phillips' ultimate classification of Sidgwick's moral position as 'utilitarian intuitionism'. The intuitionism referred to here is the epistemic device version only: It is my view that when it comes to Sidgwick's moral theory as it appears in *Rational Benevolence*, Phillips does not recognise the significance of the very 'moral-theoretic vs epistemic' distinction that he creates.

1.4.b.ii 'The Unfairness Objection', and the Question of Self-Evident Principles

Despite Phillips' recognition of Sidgwick's intuitionistic reconciliation between common sense morality and utilitarianism, Phillips himself does actually dispute the validity of that reconciliation – and he does so on the fairly novel argument that it is Sidgwick's argument for utilitarianism that is *not* successful. In order to argue *this*, he advances a strong argument which he terms 'the unfairness objection'³²³. The

³²³ Ibid. p.100. In an astute discussion of Sidgwick's moral epistemology, Phillips observes an anomaly in Sidgwick's approach to justification, which he refers to as 'the puzzle in Sidgwick's moral epistemology' (p. 65). This epistemological puzzle is relevant to my argument for Sidgwick's reconciliation because it leads Phillips to a conclusion on the type of argument that Sidgwick was using in order to reconcile the conflict between utilitarianism and Intuitionism, with which I disagree – and this point is relevant in the main text below, and later in the thesis. However, to articulate this fully would require more space than is available here, and so I only discuss this point tentatively, as far as my point below requires. Phillips remarks that Sidgwick sees two methods of argument; one Phillips calls 'criterial', which involves beginning with 'whatever we affirm to be self-evident' and then going on to seek to establish universality and originality (p. 64), and another he calls 'bipartite', consisting of course of two aspects, in which it must firstly be proven to the holder of a belief that that belief is wrong, and secondly agreed that that belief does have *some* value, in some sense. Phillips believes that Sidgwick thinks criterial arguments are stronger, and that there is a clear criterial argument in Book III, for utilitarianism, against common sense morality (p. 67). Phillips' puzzle is that having built the case that Sidgwick values – and uses – criterial arguments, he observes that they then disappear from Book IV: II, and are replaced by bipartite arguments for utilitarianism only. What, Phillips asks, is the reason for this? (p. 69). He considers various responses. He cites Schneewind as having considered Book III itself to only consist of bipartite arguments, and therefore as denying that there is an inconsistency between the two books. Phillips does not believe that this is plausible, and I agree with him in as far as Sidgwick at least *began* to develop a criterial sort of argument. It could also be argued, Phillips goes on, that there *is* actually a criterial argument in Book IV, but this is not the case either, because the criterial passages in Book IV do not appear as independent arguments, but rather only as part of bipartite arguments (p. 70), and this is not what Sidgwick considered to be a truly criterial argument (see also Sidgwick, 'The Establishment of Ethical First Principles', in, *Mind* Vol.4 (1879), pp.106-11). Phillips' own possible answer to the puzzle is that Sidgwick recognises in IV:II that the independent criterial argument against utilitarianism has failed, therefore he does not use it in Book IV, and turns instead to the other option, of the bipartite argument (*Sidgwickian Ethics*, p.72). I argue however that there is actually another answer to the apparent inconsistency, that Phillips hasn't considered. This option falls into Phillips' solution category of recognising that there is a discrepancy between the types of argument, but claiming that it is *not* a puzzle. It is that the argument in Book IV is actually a different argument - a *new* bipartite argument. There are two aspects to understanding this line of argument. The first is that Phillips has in mind at all times that Sidgwick is aiming to *establish utilitarianism against common sense morality* (pp. 67, 68, 71, 72). I argue that this is not the case. Instead, I argue that the criterial argument disappears from Book IV because a) Sidgwick had seen a connection between utilitarianism and common sense morality, and, b) this connection emerged on

unfairness objection can be summarized quite simply. In his chapter 'Utilitarianism Versus Dogmatic Intuitionism', Phillips outlines the criterial argument through which Sidgwick rejects as self-evident the maxims of common sense morality, in favour in of his philosophical intuitions (Justice, Prudence, and Rational Benevolence) that he does argue to be genuinely self-evident³²⁴. This involves Sidgwick arguing that the former do not satisfy his four criteria for self-evidence, and that the latter do. But here, Phillips says, a contradiction appears directly out of the text: Whereas Sidgwick previously argued that Intuitionism's principles be made precise enough to give 'determinate verdicts' in each case, he does not subject his own self-evident principles – which make up his utilitarian position- to the same criteria. In fact, Phillips points out, Sidgwick expressly states that his 'favoured principles' are still 'of too abstract a nature, and too universal in their scope...particular duties have still to

the basis that the utilitarian principle of benevolence could only be derived from certain self-evident premises belonging mainly to Intuitionism. Although Sidgwick finds most of dogmatic intuitionism, or common sense morality, to *not* be self-evident, he does not abandon that mode of ethics, and instead pursues it until he identifies Justice, Prudence, and Universal Benevolence – which allows him to then state the utilitarian principle. Therefore, only a bipartite type of argument – understood in Phillips' terms as first stating that a belief is wrong and then returning to the value of a belief - is relevant for this new situation in which utilitarianism and common sense morality are seen to be inextricably linked. This requires reading the passage from IV:II that Phillips cites not as a negative proof of the sudden, unexplained introduction of a bipartite argument, but rather as a positive affirmation of the fact that Sidgwick had seen the necessity of this connection/argument. This is why I say above that Sidgwick *began* to form a criterial argument for utilitarianism in Book III – but it is my view that he does not actually develop that criterial argument. "Perhaps we may say", says Sidgwick, being quoted by Phillips, 'that what is needed is a line of argument that allows the validity, to a certain extent, of the maxims already accepted, and on the other hand shows them to be not absolutely valid, but needing to be controlled and completed by some more comprehensive principle'. This is clearly a 'bipartite' argument as Phillips defines it, but it does not have to succeed where the criterial argument failed – because that 'failure' was actually not a failure, but the discovery of a new theory. It is, if we may say, an *independent* bipartite argument. Or, on a more stringent understanding of bipartite, it is not bipartite at all, but rather, perhaps, even a new criterial argument (for Sidgwick's particular type of utilitarianism, as represented by Rational Benevolence). This discussion of bipartite and criterial arguments is directly reflected in my interpretation of Sidgwick's 'negative' and 'positive' versions of proof for utilitarianism, which I lay out in chapter 3. The 'negative' proof is the criterial argument, and the 'positive' the bipartite; and I argue in chapter 3 that Sidgwick does *not* actually give fully negative proof for utilitarianism in Book III, and that he actually gives another version of 'positive' proof.

Antony Skelton is also of the opinion that the disappearance of the criterial argument from Book IV is not actually the 'puzzle' that Phillips holds it to be (see Skelton, 'Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips', a paper given at a 2013 Symposium on David Phillips' *Sidgwickian Ethics* - Anthony Skelton, « Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips », *Revue d'études Benthamiennes* [En Ligne], 12 | 2013, mis en ligne le 10 Décembre 2013 URL : <http://etudes-benthamiennes.revues.org/675> date accessed 29th June 2016. Skelton's argument for this is that Sidgwick recognized that the criterial argument would not on its own be enough to establish utilitarianism as a method of moral reasoning.

³²⁴ *Sidgwickian Ethics*, pp.96-101

be determined by some other method'. This is the unfairness objection –and on Phillips' view Sidgwick's criterial argument for utilitarianism fails because of it³²⁵.

Phillips is not alone in his view of Sidgwick's treatment of his own self-evident principles. Since Phillips, Thomas Hurka has contributed significantly to the unfairness objection with the argument that it is only if the self-evident principles are stated as 'other things equal' that they seem to meet the criteria, in which case they simply cease to be utilitarian³²⁶. Both writers observe that attempting to rescue Sidgwick's axioms requires modifications that either cause them to fail one of the other criteria, or to take on an entirely different nature from that which Sidgwick actually intended – and both point to the fact that utilitarianism most certainly fails the universal acceptance test, on anyone's reckoning³²⁷.

Together, Phillips and Hurka present a strong case against Sidgwick's argument for utilitarianism, and it has had the rare effect of challenging this usually highly

³²⁵ As Phillips says, "Sidgwick insists that common sense principles meet a standard of determinacy from which he exempts his own favoured utilitarian principles'. And that is unfair' (Ibid. p.101). The unfairness objection is quite a new point to have been brought against *ME*, given Sidgwick's long-running reputation for being nothing but methodical and systematic. It is not entirely original to Phillips; Roger Crisp, for example, also noticed the problem ('Sidgwick and the Boundaries of Intuitionism').

³²⁶ Hurka, 'Sidgwick on Consequentialism and Deontology', p. 149. Hurka's 'other things equal' argument is discussed below. Hurka also agrees that Sidgwick's own preferred moral theory falls prey to the same problems as he thought deontology to do, but he does not seem to see this, which is why it is unfair (see pp. 12, 13, 19 for the clearest outline of this). Hurka's charge against Sidgwick that his defence of utilitarianism against common sense morality is unfair turns on the same basic premise as Phillips', that Sidgwick's treatment of his own 'self-evident' principles is, in comparison, 'extremely lax' (p. 148). Hurka sees Sidgwick as having simply assumed that they satisfy the conditions, while leaving the actual substantiation of these claims extremely thin. Hurka also takes issue with the formulation of the criteria itself, arguing that if the second condition, 'the self-evidence of the proposition must be ascertained by careful reflection' (III:XI:339) can be satisfied, then surely the others are somewhat redundant (p. 131), although Hurka does vindicate Sidgwick to some extent here by positing that Sidgwick was simply careless in making this proposition, and that what he was actually referring to was the *apparent* self-evidence of a principle, as distinct from the 'mere impressions or impulses...or opinions' which with we frequently confuse genuine intuitions. Hurka also, rightly, points out that this entire process of getting an apparent intuition to meet the criteria for self-evidence in any case negates Sidgwick's description of intuitions as the result of 'immediate and irresistible' (III:II:229) cognition (p. 133).

³²⁷ *Sidgwickian Ethics*, p.105; 'Sidgwick on Consequentialism and Deontology', p.149. Phillips also presents an argument that Sidgwick's practical reason is indeterminate in 'Sidgwick, Dualism, and Indeterminacy in Practical Reason', in, *History of Philosophy Quarterly*, Vol.15, No.1 (1998), 57-78). Hurka's overall point is that it is only if Sidgwick's axioms are moderated to become 'other things equal' principles that they appear self-evident and generally accepted (and even then there are objections) (pp. 162-4), but then doing this means that they cannot fully establish utilitarianism to the degree that Sidgwick intended (they would however, presumably become Rossian-style *prima facie* principles. See below). This can only happen if they are read as 'all things considered', at which point the axioms promptly fail the condition of being widely accepted. 'As he said of deontological principles', says Hurka, 'Sidgwick's axioms may satisfy some conditions for self-evidence in one form and others in a different form, but in no form do they satisfy them all' (p. 149)

acclaimed doctrine, from this most eminent of the ‘classical’ utilitarians³²⁸. Moreover, it is clear that if Sidgwick’s self-evident axioms can be thus disputed, then this will be problematic for any theory that relies on those axioms as constructs of Sidgwick’s normative moral theory – such as my one in this thesis. Of even greater consequence for my argument is Phillips’ point that if Sidgwick has not truly established these principles – and the resultant utilitarian principle - as self-evident (and ultimately rational), then the *reconciliation* which depends on the ultimate rationality of these principles must also fail. We must therefore find a means of handling Phillips’ point³²⁹. This requires two defences; finding a line of argument that explains why Sidgwick apparently unfairly allows indeterminacy in his own self-evident principles, and addressing the more philosophically problematic issue of whether or not those principles are actually self-evident (and indeed whether this matters).

My response to the unfairness objection stands between the positions of Phillips/Hurka, who believe that it *is* unfair, and Singer and Lazari Radek, who believe that it is not³³⁰. Firstly, I acknowledge that this particular inconsistency in *ME*

³²⁸ As Phillips himself says, only C.D. Broad had made this kind of critique of Sidgwick previously.

³²⁹ Phillips does attempt to rescue Sidgwick from this charge, first by arguing that Sidgwick’s axioms actually do meet the determinacy requirement (which even Sidgwick seems to admit they do not), then by considering whether Sidgwick’s axioms meet the clarity and precision requirement when they are taken together. But again, Phillips points out, Sidgwick’s own investigations into utilitarianism do not support the idea that its maxims are clear and precise (p. 103). Phillips has also already pointed to the fact that Sidgwick ‘presents utilitarianism as itself a deduction from self-evident principles’ (p. 64). In any case, Utilitarianism does not, clearly, pass the universal acceptance test. Also, in this case, where common sense morality was not allowed to use any other principles, on Phillips’ view Sidgwick has a) combined principles or b) roped in other principles in order to satisfy the conditions. Hence, Phillips says, Sidgwick’s argument for utilitarianism against common sense morality is truly unfair.

³³⁰ *Point of View of the Universe*, p.144. Singer and Lazari Radek, focusing on Phillips’ emphasis on determinacy in Sidgwick’s criteria, argue that the problem is neither inconsistency nor leniency - and predictably do not agree with Phillips and Hurka that Sidgwick is guilty of unfairness when it comes to the testing of his own principles. Instead, they draw the reader’s attention to the difference in type and role between common sense principles and utilitarian ones, arguing that as common sense morality is specifically the collection of people’s general, intuited moral rules that *rely on common consent*, it is important that they are tested for self-evidence, in order that everyone can consent to them. Sidgwick points out that even if all our common sense morality fitted together with no conflict between them, the result would still be an ‘accidental aggregate of random precepts that are in need of rational justification’. But the morality of common sense, being as it is, does not offer that justification. This is why it relies on absolute common consent; without it common sense morality simply breaks down, and that whole system of ethics disintegrates. The dictates of a utilitarian system on the other hand, Singer and Lazari-Radek argue, do *not* themselves rely on common consent. As a system of ethics utilitarianism should still be established on a self-evident basis, but this does not mean that those principles themselves have to yield determinate results. For this reason, common sense morality must be subjected to testing for self-evidence, where utilitarianism need not be. See also Katarzyna Lazari-Radek, ‘Sidgwickian Ethics’ (Review), *Mind*, Vol.123, No. 491 (2014), 951-956. This is a nice line of defence from Singer and Lazari-Radek, and I agree with this interpretation of the different roles of

does seem to exist. Even if Sidgwick did assume that his own self evident principles safely passed the criteria, it would have been expedient of him to demonstrate – even briefly – how this is the case³³¹. But, similarly to Singer and Lazari-Radek, I do believe that there is a justified reason for the apparent anomaly. This is that Sidgwick is aware that these principles do not ‘meet a standard of determinacy’ - and that this is itself actually integral to his philosophic conclusions³³². Phillips is right that the criterial argument for utilitarianism fails. He is even right that the criterial argument fails because those principles cannot be established as self-evident (I will return to this point shortly). But, if we see the argument for utilitarianism as a *bipartite* one, and not as a criterial one (as I suggested in the notes above) it doesn’t matter that the criterial argument fails because we can argue that each of these self-evident principles does in fact rely to some extent on the others³³³. This is what makes it a ‘bipartite’ argument, and it is a *different* bipartite to the one Phillips reads Sidgwick as using *exclusively for* utilitarianism as against the type of Intuitionism that informs common sense morality. It seems to me that this is what Sidgwick is saying when he makes the point that these principles are ‘too abstract in nature, too universal in their scope’ etc. (III:XIII:379). Phillips reads this passage as Sidgwick having, ‘dropped the ball’ in terms of awareness that his own principles are not determinate. I read it as a fundamental part of his wider argument, in that he is actually *admitting* that they are not entirely self-evident³³⁴. That is, not only does Sidgwick recognise the lack of self-

common sense morality and utilitarianism. The point also leads them to their wider argument, in which they pick up Sidgwick’s statement that even under his own axioms ‘particular duties still have to be determined by some other method’ and argue that this other method is utilitarianism. This passage is also vital to my own defence of Sidgwick against Phillips, but on slightly different grounds. See the main text.

³³¹ It is this sort of apparent carelessness that leads Hurka to accuse Sidgwick of ‘equivocating on his principles’ (‘Sidgwick on Consequentialism and Deontology’, p.150).

³³² (See *Sidgwickian Ethics*, p.101). I differ with Singer and Lazari-Radek here on the point that whereas they believe Sidgwick’s system still ultimately (and safely) produces a viable form of utilitarianism, even if it is not self-evident, I believe that the failure of the criterial argument precisely does *not* produce utilitarianism, and that this is not a problem.

³³³ To reiterate, Phillips’ definition of bipartite arguments is that ‘they involve both a negative and a positive verdict on the interlocutor’s initial belief. On the one hand, negatively, the principle the interlocutor initially holds to be self-evident is not itself self-evident; on the other hand, positively, there is something right about the interlocutor’s initial principle. That initial principle is either a truth, though not itself self-evident; or, at least, though false, it points towards a self-evident truth that will emerge when its arbitrary and unjustifiable limitations are removed’ (Ibid, p.63).

³³⁴ This response also meets Hurka’s objections in the same way, given that Hurka also believed Sidgwick to be concerned only to establish utilitarianism as the supreme moral principle. To expand further, it seems likely that Sidgwick was anticipating the need for there to be more than one principle involved in establishing a fully comprehensive ethical system, and there is nothing to say that this *comprehensive* system has to itself be self-evident, especially if we borrow Singer and Lazari-Radek’s

evidence, but he understands it to be epistemologically important - and this also allows me to maintain my argument that Sidgwick was not simply defending utilitarianism.

Now for the second problem, concerning the question of whether Sidgwick's self-evident principles of Justice and Prudence can actually be established as having any degree of self-evidence. As much as Sidgwick's recognition that they are not fully determinate is important, self-evidence does matter, because self-evidence for Sidgwick is also what is ultimately rational - and what is ultimately rational for Sidgwick is also what is morally *right*, which makes it an unconditional duty. If these principles cannot be established as ultimately rational, and therefore as right without reference to ulterior results, then this threatens my entire argument that Sidgwick only *derived* the utilitarian principle from a non-consequentialist, deontological version of duty, which makes up one part of the reconciliation between deontological and utilitarian principles. It could be posited that Phillips is simply wrong that the principles are not self-evident (or that Sidgwick was just 'extremely lax', as Hurka would say, about testing his philosophical intuitions) but both Hurka and Phillips argue that if we ourselves test them on Sidgwick's behalf, they do seem to fail – and this, I admit, is a stronger point. However, this is only testing those principles against Sidgwick's *own* criteria, and it may well be that they fail this particular test. But this might not necessarily mean that their self evidence has to be rejected entirely. And ultimately, it is still more than possible to argue that *Sidgwick* believed these

line and argue that common sense morality *on its own* might depend on self-evidence, but a system that included utilitarianism also would not itself depend on self-evidence to the same extent. Phillips's unfairness objection also incorporates this very point – that Sidgwick allows the addition of further principles in order to make his own axioms pass the test where he did not allow common sense morality the same liberty. I maintain that this addition of further principles was not Sidgwick merely searching for further support for his 'own favoured principles', nor a blind sort of bias, but rather the result of these principles appearing to Sidgwick as a necessary and unavoidable stage in the exploration and explanation of the reality of human morality. Further support for the argument that the failure of Sidgwick's self-evident axioms to meet that criteria is actually indicative of his wider argument for the relationship between deontological and utilitarian principles can be taken from the personal document (included in the preface to the sixth edition). Sidgwick says here, "...I had myself become, as I had to admit to myself, an Intuitionist to a certain extent. For the supreme rule of aiming at the general happiness, as I had come to see, must rest on a fundamental moral intuition, if I was to recognise it as binding at all....and there was also the Kantian principle, which I recognised as irresistibly valid, though not adequate to give complete guidance' (xix). Over the next two pages, Sidgwick continues to explain how he oscillated between adherence to utilitarian principles and adherence to deontological ones, until eventually he concluded that there was a harmony between them (xx). Such passages, and others like them (such as the connection passages to which I refer in chapters 3 and 4) significantly support the idea that, for Sidgwick, neither method of ethics was fully adequate on its own.

principles to be self-evident, and therefore that for him, the derived principle of Rational Benevolence *is* ultimately rational – and thus *is* unconditionally right, without reference to consequences. Overall, I anticipate that a combination of these responses to Phillips (and Hurka's) quite legitimate claims against Sidgwick on this issue is sufficient enough to still allow for my particular construction of Rational Benevolence, and for that theory to hold up philosophically³³⁵.

1.4.c. Thomas Hurka: Sidgwick's 'Extreme' Deontology

With Phillips, Sidgwick's utilitarian status has received the curious treatment of having been both maintained – and questioned for its integrity. Thomas Hurka makes a similarly unusual move, by being one of the very few theorists to specifically discuss Sidgwick's deontology. Even more surprisingly, Hurka interprets this to be an 'extreme' understanding of deontology, which contrasts unfavourably with the later 'moderate' deontology of W.D. Ross³³⁶.

³³⁵ I do, however, point to a possible need for future work on the self-evidence of Sidgwick's principles. Here, Brad Hooker offers something extremely valuable, and something that could possibly be a far more effective way of expressing 'self-evidence' as it appears in the context of this thesis. Hooker is wary of the very term 'self-evident', believing it to have so often 'been misused by popular moralists to express an unreasonable degree of certainty' (Hooker, 'Intuitions and Moral Theorizing', in, Stratton Lake, *Ethical Intuitionism*, 161-183, p.165). Instead, Hooker prefers the term 'independently credible', which refers to a belief that is 'attractive without reference to evidence beyond itself and yet might turn out to be mistaken'. Both of these statements from Hooker accord closely with what would seem to be the most appropriate conclusion to draw on Sidgwick's self-evident principles. It certainly appears from the strength of the Unfairness Objection that Sidgwick himself was at least guilty of using the term 'self-evident' to express that 'unreasonable degree of certainty' to which Hooker refers, and yet Sidgwick's axioms do attain a [degree] of self-evidence that would seem to justify them in at least being called 'independently credible'. This is the line I take overall, against Phillips' argument for unfairness.

Phillips does also take note of this point from Hooker, and briefly considers to what extent Sidgwick himself made a distinction between strong and weak versions of self-evidence (*Sidgwickian Ethics* p. 86). He concludes that Sidgwick was aware of apparent self-evidence, but only seemed to believe that this meant additional external support was required, rather than considering the idea that some evidence for self-evidence could be supplied by contemplation of the principle itself. Overall, Phillips agrees that Sidgwick's case can be supported more effectively if a weaker form of self-evidence is allowed for, and argues that this is actually present in Sidgwick's work (p. 59)

³³⁶ Hurka is also the editor of *Underivative Duty: British Moral Philosophers from Sidgwick and Ewing* (Oxford: Oxford University Press: 2011), and *British Ethical Theorists from Sidgwick to Ewing* (Oxford: Oxford University Press: 2014), which also include examinations of the common ground shared by consequentialists and deontologists such as Sidgwick and Ross. Given that his 2014 paper, 'Sidgwick on Consequentialism and Deontology: A Critique' is the most relevant to my argument here, this will be the main text used in this discussion of Hurka's reading of Sidgwick. Where relevant (and possible) however, *British Ethical Theorists*, and *British Moral Philosophers*, will be used for their additional points.

In 'Sidgwick on Consequentialism and Deontology; A Critique', Hurka begins by arguing that Sidgwick argued for consequentialism, and against deontology. Sidgwick's critique of deontology, Hurka outlines, involved arguing that where some of the principles might seem to have a degree of universal acceptance, any qualifications that are added to them immediately destroys that consensus, and thus blocks their self-evident status³³⁷. Hurka then goes on to say that the pluralist deontology view (with which Sidgwick was working) requires that we weigh conflicting duties, but gives us no means for doing so. Overall, Hurka points to Sidgwick's problems with the indeterminacy of common sense morality principles, and the lack of clarity, certainty, and precision inherent in this method. And this, Hurka argues, is the direct result of Sidgwick having failed to grasp 'the concept of a prima facie duty and therefore not seriously considering the moderate, as against absolute, deontology that it makes possible'³³⁸. That is, Hurka believes that Sidgwick should have taken account of an 'other things equal' version of a deontological system, in which principles are not always inviolable. Instead, Sidgwick seems to take the 'right all things considered' line, which only allows for absolute, or extreme, deontologies³³⁹. Hurka summarises, 'he in effect assumed that the only alternative to consequentialism is an absolute deontology which forbids some acts whatever their consequences'³⁴⁰.

Hurka's offensive on Sidgwick's failure to see the value of prima facie duties is convincing, turning on the idea that what Sidgwick thought to be derived from the principle of utility was actually a form of moderate deontology, readily available to him, which he simply missed and which caused the grounds for his objections³⁴¹. In

³³⁷ 'Sidgwick on Consequentialism and Deontology', pp.133-134

³³⁸ Ibid. pp.135, 148

³³⁹ Ibid. p. 136. To be exact, Hurka disputes what Sidgwick meant when he defined the Intuitionist position as holding that we have the capacity to see that certain acts are 'right and reasonable in themselves'. If, Hurka argues, by this Sidgwick meant that they are seen as right 'all other things equal', then Sidgwick had in fact grasped the idea of a moderate deontology. If, however, he meant right 'all things considered', then this is an extreme form only.

³⁴⁰ Ibid, p.136

³⁴¹ Ibid, p.9-12. He also further challenges Sidgwick's position by pointing to the fact that Sidgwick's objections to the deontological position apply just as much to his own consequentialist principles. To cite a few examples, Hurka does admit that Sidgwick had a point that if a moral truth is self-evident, we should be able to intuit it with precise content ('Sidgwick on Consequentialism and Deontology', p.143) - but does this apply to just deontological theories, or to all claims to self-evident moral knowledge? If so, Hurka points out, it cannot support consequentialism over deontology. Hurka refers to Sidgwick's principle of Justice as his 'universalizability' axiom (p.144), but makes the point that this is not distinctly consequentialist and that most deontologists accept it (which, I argue, is actually evidence for Sidgwick's recognition of certain deontological values). He also takes a strikingly

light of Hurka's argument that Sidgwick's system would have benefitted from his seeing the possibility of *prima facie*, rather than absolute, duties, the question arises as to whether *this* would be a more effective way of tackling the boundaries between deontological and utilitarian moral properties, as opposed to the sort of reconciliation between deontological and utilitarian principles for which I am arguing here³⁴².

different view of the point of view of the universe to Singer and Lazari-Radek, by stating that the point of view of the universe – which only appears twice in *ME* and then tentatively - is only true if one *does* accept that view. Sidgwick never said 'you ought to adopt the point of view of the universe', which means that as an axiom it still does not give any actual practical guidance – which of course is one of the objections Sidgwick brings against deontological principles (p.146). Overall, Hurka raises many issues with the credibility of Sidgwick's utilitarianism, and of the principle of Benevolence – but those are not my concern here. My focus in this thesis is to give metaphysical grounds for there being a mutually informing connection between deontological and utilitarian properties of morality, and to use Sidgwick to do so – the issue of whether Rational Benevolence is actually as sound as Sidgwick believed it to be is an important one, but for now it can be detached from the example Sidgwick offers of how the two types of moral principle combine to work together. Having said this, I do want to state my belief that there are grounds on which to argue that Sidgwick did actually see the point of view of the universe as a specific moral duty. He uses the phrase 'as a rational being I am bound to aim at good' (III:XIII:382), and again a few lines later, 'morally bound'. We know from Sidgwick's earlier discussion of 'right' and Reason that rational beings *ought* to pursue those things that they recognise as an ultimately rational end. The combination of rational beings and rational ends is what therefore what incurs the ought that Hurka thinks is missing.

³⁴² Hurka makes a good case for Ross's being able to answer most of Sidgwick's objections. Hurka uses Ross to demonstrate how recognition of *prima facie* duties, and 'other things equal', avoids the need for criteria for self-evidence like the one Sidgwick develops. Applying Ross retrospectively to Sidgwick, Hurka argues that if Sidgwick had recognised this, he would have avoided many of the problems of trying to establish self-evidence, which in the end were self-defeating. Hurka points out that Ross is able to describe how two duties may be right at once, while seeming to conflict. They do not contradict each other, it is simply that on the basis of other things equal it is possible to devise which duty should be followed, without the other duty being wrong, or ceasing to be a duty. Presented in this way, Ross' system does seem to be able to deal with conflict between principles more effectively and simply than Sidgwick, who – on Hurka's interpretation - stated that in a situation of conflict, one or the other principle, or both, must be wrong. This then requires Sidgwick to introduce exception-clauses, which of course causes the principles to fail self-evidence criteria 2 and 4 ('Sidgwick on Consequentialism and Deontology', p. 137). (Phillips also draws a comparison between Sidgwick and Ross. There is, Phillips suggests, an 'unofficial' way of arguing for utilitarianism, which does not involve common sense morality having to generate determinacy. This involves recognising only *apparent* self-evidence. Here, Phillips argues, Sidgwick's main opponent is Ross, whose 'deontological intuitionism explicitly and by design does not meet the clarity and precision condition interpreted as requiring complete determinacy' (*Sidgwickian Ethics*, p. 105) Where Sidgwick argues that principles such as 'I ought to keep my promises' are not self evident, Ross believes that they are (although this is *prima facie* self-evidence, and may in fact only be metaphysical, and not epistemic). So, Phillips says, if even a deontologist's principles cannot be rejected on the grounds of indeterminacy, how can Sidgwick be justified in rejecting them?)

I do, however, think that Hurka himself is guilty of some unfairness towards Sidgwick when it comes to his comparison between Sidgwick and Ross. Hurka drags Sidgwick's search for self-evident principles, and his subsequent failure, over the coals, but the theory to which he is unfavourably comparing Sidgwick's i.e. Ross's does not itself rely on the same sort of self-evidence that Sidgwick was aiming to establish. This is because Ross has no *need* for self-evident principles, as Phillips also points out in his similar comparison. Ross's system is a weighing of duties, the result of which are never self-evident (*Sidgwickian Ethics*, p. 105) Ross does not 'solve' the problem of self-evidence therefore, he avoids it. Hurka himself uses the language of avoidance, rather than solution (arguing that if Sidgwick had recognised the value of *prima facie* duties he could have avoided the self-defeating problems of trying to establish self-evidence) but the problem remains; comparing a theory that fails to do something to another theory that does not even attempt that same thing hardly seems

Hurka's exposition of how Ross' system may have an advantage over Sidgwick's is, I believe, among the best – and presented in this way the *prima facie* duty approach becomes a viable candidate for reconciliation between deontological and consequentialist considerations. However, I take a different and more positive view of Sidgwick's use of 'extreme deontology'. As I outlined in the thesis introduction (and as will be continued properly in the following chapter) it is indeed Hurka's 'all things considered' version of deontology that Sidgwick uses, and furthermore it is *this* deontological view i.e. an absolute, non-consequentialist one, that I argue informs his own concept of rational moral obligation. This in turn provides the crucial traditionally non-consequentialist deontological basis of Rational Benevolence, which I then argue unites with his utilitarian principle to create a mutually informing synthesis between equally rational deontological and utilitarian moral properties. This of course requires arguing, contra Hurka, that it is an advantage to Sidgwick's system that he takes the extreme view of deontology, which I essentially do in the next chapter³⁴³. But, following on from the construction of Rational Benevolence, this is not where I believe Sidgwick's version of deontology to end. There is scope to argue that Rational Benevolence itself represents a form of 'moderate deontology', in that the addition of the utilitarian principle to a foundation of non-consequentialist duty adds a *prima facie*, 'other things equal' dimension to the moral theory as a whole. This tempers Sidgwick's original 'right all things considered' interpretation, and may actually contribute helpfully to the idea of a synthesis between consequentialist and non-consequentialist moral principles³⁴⁴. Taken altogether then,

fair. Hurka seems to play Sidgwick at his own game, in critiquing his search for self-evident principles, but then attempts to win according to the rules of a different game. Furthermore, where Ross does talk about self-evident principles, his claim is that self-evident moral truths may well exist metaphysically but are epistemically unavailable to human beings, limited as we are in time, space, and knowledge. Sidgwick's epistemology attempts to go further than Ross's here (with his testing of intuitions for ultimate self-evidence) which creates further distance between the two intuitionists.

³⁴³ In chapter 2 I do not make the specific point, in Hurka's terms, that Sidgwick is using the 'all things considered' version of deontology, as opposed to the 'other things equal' – but I can clarify here that it is the 'all things considered' version that I am assuming of Sidgwick's deontology, as it appears before it becomes part of Rational Benevolence, where it becomes more moderate (see main text above).

³⁴⁴ This is not to say that Ross's system would simply be superior, as Hurka is suggesting here. I theorise Sidgwick's Rational Benevolence to be *similar* to Ross's argument for *prima facie* duties, in that they both hold neither principle in situations of moral conflict to be actually *wrong*, even if the dominance of one is temporarily recognised for the purposes of a given situation. As I outlined in the introduction (with reference to Hare, who did expressly believe that in a situation of conflict of duty, 'one of them is not your duty'), the way in which Sidgwick constructs Rational Benevolence unites the fundamental, non-consequentialist duty-based moral property that underpins the morality of common sense with an equally fundamental utilitarian-based moral property, and *this* can be used to overcome conflicts between duties. There is, in any case, perhaps scope to engage further with Hurka on this

Sidgwick may have *begun* from an extreme deontological position – and this may be crucial to the synthesis – but it is *not* the system that he himself eventually offers in his own moral theory. This system, I theorise, is much closer to a version of moderate deontology.

Hurka's wider work raises another issue about how duty features in Sidgwick's work, which it is extremely important I take account of here, ahead of continuing my own argument for Sidgwick's particular notion of duty. One of the points central to Hurka's two most recent books is that underivative duty, i.e. the notion that basic moral duties are neither reducible to nor derivable from non-normative judgements, appears as much in Sidgwick's consequentialism as it does in Prichard and Ross's deontology³⁴⁵. Hurka summarises, 'if asked whether we ought to promote other people's happiness, they would have said that there is no answer other than we ought to...no less than the deontological duties of Prichard, Ross, and Broad, the supreme consequentialist duty holds just because it does'³⁴⁶. In Hurka's concept of underivative duty then, we see clearly the attributing of a strong sense of *duty* to Sidgwick that is actually not connected with the traditional 'deontological' school, and that pertains wholly to his consequentialism. Hurka furthers the distinction between 'duty' as it appears in each school of thought by pointing out that for Sidgwick, the fundamental duty is to promote the good, as opposed to deontologists such as Ross, for whom deontological duties cannot be derived from this more fundamental duty to promote the good³⁴⁷. I entirely support Hurka's categorisation of this as a 'thin' concept in Sidgwick; as Hurka cites, Sidgwick believed that 'ought' was 'too elementary to admit of any formal definition'. Sidgwick also most certainly argues for a type of duty that is in part consequentialist – and Hurka's view that Sidgwick considered the most fundamental duty to be the promotion of good is admittedly a difficult point to challenge. But on my view this is not the *only* notion of duty that

point about Sidgwick and prima facie duties, and to ask whether Sidgwick might have actually recognised the *concept* of prima facie duties to some extent even if he did not expressly develop it actually in the pages of *ME*.

³⁴⁵ Hurka, *Underivative Duty*, pp.24-25; *British Ethical Theorists*, p.1. This underivative duty, Hurka argues, is a fundamental common denominator between these positions –even more so, perhaps, than intuitionism.

³⁴⁶ *Underivative Duty*, p.25.

³⁴⁷ Hurka also characterises Sidgwick's irreducible moral concept (or his 'radically minimalist view') as being 'what a person ought, has a duty, or as some today would say, has a reason to do' (*British Ethical Theorists*, p.22).

Sidgwick uses. As I outlined in the thesis introduction, and will argue in full in chapter 2, I believe that Sidgwick does in fact also maintain a traditionally deontological notion of duty, that is Kantian, and *non-consequentialist*, and that *this* notion of duty does inform the principles from which he derives utilitarianism. On the basis of the fact that I believe Sidgwick's moral theory to be an amalgamation of non-consequentialism and consequentialism however, the ultimate moral duty in Sidgwick – that of Rational Benevolence – emerges as *both* deontological and utilitarian in nature.

1.4.d. Roger Crisp: Sidgwick's Deontology, and the Lack of Rational Synthesis

Roger Crisp – and his excellent book *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics* - provides a neat chronological and conceptual ending to the list of theorists included in this narrative³⁴⁸. His comprehensive exegetical and critical engagement with *ME* provides a model of Sidgwick's place in moral philosophy today –and although there is not space here to engage with the whole volume, he collates many of the themes that have occupied this chapter and discusses some highly important topics in terms that are immediately relevant to much of what has been said above. The main example of this is his discussion of Sidgwick's treatment of deontology, which accords with the interest in this aspect of Sidgwick's work revitalised by Hurka. He also makes several more points that align closely with my own theories about Sidgwick – and one which disagrees significantly - thereby providing a convenient point from which to go on and develop those claims in the rest of the thesis.

I will begin with the more general points. Firstly, Crisp actually begins *Cosmos of Duty* with the fact that Sidgwick's plan for *ME* was that it would 'reconcile 'moral sense', or 'intuitionism', with utilitarianism'. He immediately follows this with the claim that whereas Sidgwick is usually described as the third of the great 'classical utilitarians', his own personal account of his intellectual journey – as seen in the preface to the sixth edition – shows that he soon departed from his initial adherence to utilitarianism as strictly understood, 'never to return'³⁴⁹. Crisp also points to the

³⁴⁸ Crisp, *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics* (Oxford: Oxford University Press: 2015)

³⁴⁹ Ibid. p.5

impartial position that Sidgwick intended to adopt for *ME*, and the fact that he was nonetheless led to a moral theory. All of these points are identical to those from which my own use of Sidgwick begins (except of course, the last, where Crisp considers this moral theory to be exclusively utilitarian, where I do not).

Crisp also draws attention to Sidgwick's denial of rightness as means to ends (and irrespective of good), and describes Sidgwick's criticism of the 'instrumentalist view of rightness' as 'powerful'³⁵⁰. Crisp interprets Sidgwick, as I do, as working with a normative conception of rightness, in which recognition of rightness involves a 'definite precept to perform it', which results from Reason, or ultimately rationality, as the motivating factor³⁵¹. Crisp contrasts this with Sidgwick's view of goodness, which leads to a definite distinction in Crisp's reading of Sidgwick between rightness as having the only normative property, and goodness as a mere evaluative judgment³⁵². Crisp waives this to some extent later in the text, by arguing that the 'bestness' within goodness implies a rational dictate to perform the action *if one can*', but his overall position on Sidgwick's distinction between rightness and goodness maintains the idea that rightness is associated with rational dictates³⁵³. This supports my reading of Sidgwick's view of rightness, which is in turn a crucial component for my understanding of Sidgwick's argument for Rational Benevolence, and for my argument that both deontological and utilitarian properties are presented as ultimately rational by Sidgwick, and therefore as ultimately right.

Crisp offers the helpful term 'non-consequentialist intuitionism' with which to distinguish the form of intuitionism in which actions are known to be right in themselves, independent of their consequences, and also argues that Sidgwick's intuitionism represents a 'standard argument for foundationalism in epistemology'³⁵⁴.

³⁵⁰ Ibid. p.12

³⁵¹ Ibid. p.62

³⁵² As Crisp describes it, the moral ideal in goodness is to be construed as attractive, not imperative (Ibid.p.57). We could, Crisp says, make a case that goodness is actually 'non-latently normative' – but this is not the view Crisp believes Sidgwick to take, given Sidgwick's connection between rightness and normative Reason (p. 59).

Crisp does, however, question whether Sidgwick's definition of ultimately reasonable must equate with a *moral* judgement. i.e. whether what is ultimately reasonable must also be that which is morally *right*, as represented by an 'ought' (pp. 17-18). His own views, Crisp argues – of egoism and utilitarianism – 'are best stated in terms of ultimate reasonableness, with no reference to ought or morality at all' (p. 18). This is a similar position to Hurka's rejecting the point of view of the universe as being an 'ought', and it would seem as though Crisp would agree with him on this. Later however, Crisp leaves behind that distinction when he argues that Sidgwick's rightness is indeed a normative conception (p. 57. See also p. 71)

³⁵³ Ibid. p.63

³⁵⁴ Ibid. pp.99-102

Crisp does, however, see the coherentist position in what he considers to be Sidgwick's reluctance to abandon an Aristotelian approach to testing philosophical conclusions against common thought³⁵⁵. And here, within the realm of that ongoing 'Sidgwick as a foundationalist/coherentist?' debate, we reach one of the most interesting readings of the *implications* of Sidgwick's apparent coherentist tendencies. For it is due to Sidgwick's allowing of a semi-coherentist approach to common sense morality that Crisp believes his argument against deontology is unsuccessful. In order to fully contextualise what Crisp says here, it is helpful to look at his relevant preliminary comment made in the preface to *Cosmos of Duty*. Here, Crisp states that there are two failures in ME. The first was the one Sidgwick perceived – that between egoism and utilitarianism, and Crisp thinks (like Phillips) that this can actually be resolved. The second failure, he argues, is 'his attempt to dismiss a reflective form of pluralistic deontology of the kind defended by Aristotle, and developed by Ross'³⁵⁶. On Crisp's view, Sidgwick is correct to promote hedonistic utilitarianism – but Crisp states that he himself cannot adhere to this position on the basis of Sidgwick's argument for it (and against common sense morality) in ME³⁵⁷. That is, Crisp makes startling use of his personal view to point to the fact that he does not believe Sidgwick to have done enough to truly *refute* common sense morality, or the deontological position. Returning now to the main text, and to how this is the result of Sidgwick's coherentist inclinations, Crisp has pointed out that the rules of common sense morality cannot, for Sidgwick, be made precise or consistent³⁵⁸. He then shows that Sidgwick did allow a degree of authority to common moral thought –but that in doing so, Sidgwick did not explain why a dogmatic intuitionist should not find that their common sense morality principles could be thought of as 'primary intuitions of reason'³⁵⁹. Why, then, Crisp asks, does Sidgwick not distance himself entirely from that source of morality? Instead, he takes the coherentist route of attempting to test common sense morality via philosophical

³⁵⁵ Ibid. p.107. Crisp also associates this non-consequentialist intuitionism directly with common sense morality, which gives a particularly clear view of Sidgwick's frequently-noted confusing grouping together of apparently very different methods of ethics under the one term 'intuitionism'.

³⁵⁶ P. vii

³⁵⁷ Crisp says, 'his own account of the epistemic implications of disagreement requires me, and indeed other moral theorists, to suspend judgement on the question of which ethical view is correct' (vii).

³⁵⁸ Ibid. p.104

³⁵⁹ Ibid. p.105

theory, and in doing so leaves it unclear how or why we are to decide between the deontological position, and his utilitarian one.

Crisp, then, has presented a view of Sidgwick's coherentism which not only does *not* offer a form of reconciliation between deontological and utilitarian principles, but which actually finds in favour of the deontological position. It is interesting that Crisp maintains deontology in *ME* in this way, but it is not the same way in which I aim to maintain it. This brings us to the point to which I alluded in the introduction to this section, that there is a point on which Crisp diverges from my otherwise often comparable reading of Sidgwick. Crisp points out that of the two problems with common sense morality that Sidgwick identifies – lack of clarity/consistency, and the lack of a single unifying principle which could provide a 'rational synthesis' – Sidgwick pays very little attention to the second problem. The result was, as was seen, that dogmatic intuitionism can still contest utilitarianism. 'Disappointingly', Crisp says, 'all we have in the *Methods* at this point is assertion. If Sidgwick could have supplied a powerful argument for rational synthesis he might have omitted from *The Methods* much of the detailed discussion of common sense morality...' ³⁶⁰. Crisp is not arguing, as Hurka did, that Sidgwick could have systemised common sense morality more effectively if he himself had adopted the pluralistic Aristotelian/Rossian deontology that Crisp thinks Sidgwick is left with ³⁶¹. Crisp is arguing precisely that Sidgwick failed to offer a synthesis at all – and that this is part of what leaves his

³⁶⁰ Ibid. p. 104

³⁶¹ Crisp refers again to the 'Aristotelian dialect' which he says Sidgwick runs 'unstably' alongside philosophical intuitionism at p.212. 'Aristotelian dialect' is a wonderful description of how Sidgwick presents many of his arguments, and it is original to Crisp in as far as applying it to Sidgwick goes. It might also be asked whether this Aristotelian dialect might in itself be another form of reconciliation, mediating as it does between common sense morality and philosophical intuitionism. To this I argue that whereas the eventual relationship between the two positions might look rather like this sort of system, this would be a by-product of the synthesis for which I am arguing, not the intended mode of synthesis itself.

This leads to a wider point, that Crisp's focus is very much on the relationship between common sense morality and utilitarianism as two methods of ethics (Ibid. pp. 198-234). His analysis of this is exceptional in depth and clarity, and I am aware of the challenge of such exegesis to my claim that the deontological and utilitarian principles that inform common sense morality and utilitarianism respectively are actually mutually informing. Furthermore, Crisp also draws on passages to show Sidgwick's 'proving' of utilitarianism, that I also use to show a *relationship* between common sense morality and utilitarianism (see pp. 209-214). But ultimately my focus is much narrower, and it is on Sidgwick's construction of Rational Benevolence as a moral theory. Whereas it is important that Sidgwick's apparent argument for *just* utilitarianism (in Book III and in Book IV of *ME*) is accounted for, I do believe there is a case to be made for the mutually informing nature of deontological and utilitarian principles in Rational Benevolence being visible in Sidgwick's exposition of the relationship between the two methods, even as it appears in that 'proof', both positive and negative. This case is developed in full from the end of chapter 3 of this thesis, and into chapter 4.

position open to a challenge from deontology. Ironically then, Crisp's argument is not that Sidgwick does not offer a synthesis between deontology and utilitarianism because Sidgwick's system was exclusively utilitarian. His view is that Sidgwick didn't offer such a synthesis because his position is actually *not* as convincingly utilitarian as is commonly assumed. I argue of course, contra Crisp, that Sidgwick's moral theory does present a synthesis between deontological and utilitarian moral principles. It is not specifically a synthesis that aims to systemise *common sense morality* and utilitarianism, although as I will argue in later chapters I do believe that the meta-ethical synthesis in Rational Benevolence maps onto how Sidgwick sees the relationship between those two moral positions. Lastly, where it is Crisp's view that Sidgwick's leaving his argument exposed to the deontology in common sense morality was accidental and disadvantageous, I will be arguing that where deontology does feature in Sidgwick's moral theory of Rational Benevolence, this is entirely intentional, and expedient to the theory.

1.5. Conclusions

Having reached the most recent work to be done on Sidgwick, this chronicle of how his work has been interpreted throughout the years can come to a close. There are of course a thousand other discussions on Sidgwick, on various aspects of his work, all of which will be valuable in some way to any reading of him; it is with great regret that so many theorists have had to be excluded from this discussion³⁶². But the aim

³⁶² I will attempt here to give a brief overview of other eminent scholars who have contributed to the now extensive range of work done on Henry Sidgwick, but whose work has not fitted directly into the direction of the narrative I have tried to construct in this chapter. Focus still predominantly remains on subjects that are particularly relevant to my arguments in this thesis.

Firstly, John Deigh, mentioned above in relation to where Sidgwick is possibly more Kantian than deontologist Ross, agrees that Sidgwick borrowed vocabulary and distinctions from Kant ('Sidgwick on Ethical Judgement', p. 244). He specifically addresses where Sidgwick's use of Reason and ought places him in terms of judgements of duty, and believes Sidgwick to be an internalist, where ought implies a reason for action, as opposed to an externalist, which does not argue that Reason has this role. David Brink has argued (in 'Sidgwick and the Rationale for Rational Egoism', in, Schultz (ed.) *Essays on Henry Sidgwick*, p. 199-240) that Sidgwick is both an internalist and an externalist, but Brink favours the externalist view, due to the Dualism of Practical Reason, which means that judgements of duty do not necessarily bring reason to action. Deigh maintains however that Sidgwick thought *moral* rather than prudential judgements to be 'the clearest examples of what he meant by an ethical judgement, and, correspondingly, that he took the notion of duty to be the primary or clearest representative of what he meant by the fundamental notion of ethics' ('Sidgwick on Ethical Judgement', p. 247) – and my view aligns with Deigh's (see chapter 2, p. 8). Deigh then goes on to point out that Sidgwick used the categorical imperative differently in relation to moral judgements of

here has been to tell a very particular story about Sidgwick, about his utilitarianism, his deontology, and about his applicability to both. Despite the length of the story, its chapters can now be easily summarised.

From the earliest days of *ME*'s publication, Sidgwick was read as a utilitarian.

Despite some cursory recognitions that this was perhaps not quite the standard utilitarianism of Bentham and Mill, time and tide did not wait for this impression.

Sidgwick then continued quietly to be as relevant to important deontologists such as Ross, Frankena, and Rawls, as he was to twentieth century revivals of utilitarianism.

More recently, deeper critical engagement with Sidgwick has furthered his utilitarian status, *and*, tellingly, drawn attention to where deontology features in his work.

Sidgwick has even been relevant to one of the greatest deontological/utilitarian

'ought' as opposed to prudential judgments of ought. He does, however, point out that Sidgwick departs from Kant on his inclusion of hypothetical imperatives (which incidentally furthers Deigh's internalist account. See p. 252), and for Deigh, it is actually a problem that Sidgwick did not distinguish between right and ought in the hypothetical and categorical imperatives. On Deigh's view, Sidgwick's use of Reason emerges as almost more Kantian than Kant's. Ross Harrison also agrees with Deigh's view that Sidgwick was an internalist ('The Sanctions of Utilitarianism', in *Proceedings of the British Academy*, Vol. 109 (2001), 93-116). See also Robert Shaver's argument that Sidgwick only holds a thin, non-reason-giving account of the good, in 'Sidgwick's False Friends' (in, *Ethics* Vol.107, No.2 (1997) 314-320).

Antony Quinton (*Utilitarian Ethics* (London: The Macmillan Press Ltd.: 1973) treats Sidgwick as a critic of utilitarianism ('rather than as a continuator of it' p. 87), and although he does ultimately define Sidgwick as a utilitarian, he outlines four counts on which Sidgwick differs from classical utilitarianism. The two that are the most important to my discussion are: that pleasure is not, for Sidgwick, the sole object of desire (p.88. See chapter 3 of this thesis for the discussion of Sidgwick's idea of pleasure), and Quinton's observation that Sidgwick's intuitionism is partly substantive, and partly *a priori* (p.89). This latter point especially accords with my own reading of Sidgwick's intuitionism, in that I believe it to be partly 'purely' intuitional, but also further verified by Sidgwick's search for self-evidence. Quinton also attaches Sidgwick's argument that a rational being ought to aim at good specifically to Rational Benevolence, and states this to be 'an improved, rationalised, version of the fundamental principle of utilitarianism' (p. 90). He also argues, as I do indirectly, that common sense principles are 'implications' of Rational Benevolence. The main difference, of course, is that I argue this Rational Benevolence is not exclusively utilitarian.

Stefan Collini ('My Roles and their Duties: Sidgwick as a Philosopher, Professor, and Public Moralist', in *Proceedings of the British Academy*, Vol. 109 (2001), 9-49), states that 'even by the standards of moral philosophers, Henry Sidgwick showed a striking readiness to use the word 'duty'. John Skorupski, who describes Sidgwick's ethical position as 'philosophical intuitionism' (*English Language Philosophy* 1750-1945 (Oxford: Oxford University Press:1993, p.67), gives an innovative account of why there is either unity in Sidgwick's practical reason, or actually *three* postulates, as opposed to the infamous dualism (this is based on Sidgwick's argument that intuitive knowledge must also be made precise, and therefore counts as much as a part of practical reason as utilitarianism and egoism). See also Skorupski's *Sidgwick and the Many-Sidedness of Ethics*, in Bucolo, Crisp and Schultz (eds.) *Henry Sidgwick: Happiness and Religion* (2004), 410-422). J.L. Mackie gives a particularly good defence of Sidgwick's dualism of practical reason, agreeing with Sidgwick that it was irreconcilable, and that if Sidgwick's approach could not reconcile it, then certainly no-one who has come after him could either ('Sidgwick's Pessimism', *Philosophical Quarterly* Vol. 26, No.105 (1976) 317-27. See also David Holley, 'Sidgwick's Problem', in, *Ethical Theory and Moral Practice*, Vol. 5, No.1 (2002), 5-65).

reconciliation theories that has perhaps ever been put forward – Parfit’s Triple Theory. Yet through all of this, his own reconciliation, as it appears in *ME*, has been missed. Theories such as coherentism and reflective equilibrium have offered worthy alternative ways of reading *ME* as offering some kind of reconciliation between deontological and utilitarian principles, but I maintain that none of these are reconciliations of the kind that I am proposing to be available in *ME*. These reconciliations have usually pertained to Sidgwick’s merging of common sense morality with utilitarianism, and whereas this is part of my theory, it is not the *meta-ethical* reconciliation between deontology and utilitarianism that I believe Sidgwick saw. The two positions are presented with utilitarianism in control - systemising, supplementing, correcting, directing common sense morality - and Sidgwick’s utilitarianism is accepted as it is. Certainly no other reconciliatory theory has applied Sidgwick’s language of ‘synthesis’ to the way in which he constructs Rational Benevolence. I suggest that the reading of Sidgwick that I offer here will itself synthesise with these other readings, in such a way that the various movements towards recognising Sidgwick as having applicability to both deontological and utilitarian positions might be thrown into a different, and possibly revealing, light.

CHAPTER 2

BUILDING A SYNTHESIS: SIDGWICK'S ACCOUNT OF DEONTOLOGY

The thesis introduction and chapter 1 were intended to together set up the thesis' premises, and the grounds for those premises. For the purposes of clarity, those ideas can be summarised as; a) frequent moral ambivalence experienced by many individuals indicates the moral value of both deontological and utilitarian principles of morality, b) the validity of the two types of principle indicates that both are (to an extent) reasonable, c) an understanding of human morality is needed which accounts for this dual reasonableness, and d) Henry Sidgwick's moral theory of Rational Benevolence, which connects the two types of principle on their shared appeal to Reason, provides this understanding. This chapter now marks the beginning of this thesis' own investigation into the research question proper, represented by that last point, of whether Sidgwick's theory of Rational Benevolence can be found to afford significant moral value to both deontological and utilitarian principles, holding them in such a way that they are no longer divided, but rather mutually inform each other in a type of relationship that could be called a 'synthesis'.

The first step in this process is to establish an irrefutable presence of deontological principles in Sidgwick's moral theory³⁶³. Whereas the utilitarian aspect of Sidgwick's work has been much expounded, it is only if it can be found that Sidgwick also attributes value to traditionally non-consequentialist deontological ideas that it becomes possible to argue for a Sidgwickian moral theory that depends on deontological principles as much as it does utilitarian ones³⁶⁴. To this end, this chapter focuses specifically on Sidgwick's interpretation of deontological moral concepts, and the extent to which he relies upon them in his own moral theory, as an integral part of human morality.

This analysis of Sidgwick's deontology follows the chronological structure of *ME* itself, and is organised into two stages. The first is concerned with Sidgwick's own understanding of deontological principles of morality, and tracks those principles

³⁶³ As outlined in the thesis introduction, 'deontological' and 'deontology' are herein defined as representing the moral view that actions are right in as far as they conform to certain moral rules, or *duties*, that are considered to be absolutely binding, and unconditionally right, without reference to consequences.

³⁶⁴ It is also through this argument for Sidgwick's deontology that the thesis is able to challenge the long-standing idea that Sidgwick is simply a utilitarian.

from their earliest manifestations in *ME*'s introduction, through to Sidgwick's argument for the connection between Reason and moral rightness. The purpose here is to clarify that by 'Intuitionism' (with a capital 'I') Sidgwick is referring to a recognisably deontological position, understood as representing non-consequentialist and unconditional duty, and then to argue that despite his utilitarian reputation Sidgwick actually holds a similar and largely non-consequentialist view of rightness – and that this contributes significantly to his own understanding of moral obligation.

The next part of the chapter then examines the value that he ascribes to deontology as it is represented by Dogmatic Intuitionism (or common sense morality), first as a method of ethics important to the ordinary individual's lived experience, and then in his own moral theory. This leads to the central discussion of Sidgwick's self-evident principles - Justice, Prudence, and Universal Benevolence – which crucially argues for the ultimately rational and non-consequentialist foundation of those principles. The aim here is to show that Sidgwick derives Universal Benevolence as a non-consequentialist moral obligation from a recognisably deontological and even Kantian basis, thus showing that part of Sidgwick's 'Intuitionism' actually informs his own epistemic 'intuitionism'. Kant's influence here is further supported via reference to the personal document (PD).

The chapter concludes that Sidgwick affords far greater value to a version of traditional deontological duty than is commonly thought, not just as a method of moral reasoning to which an ordinary person does recourse in their moral thinking, but also as an indispensable aspect of his normative morality represented by the term 'Rational Benevolence'. The chapter also concludes that despite this Sidgwick does not consider deontological principles to be by themselves sufficient for constructing a wholly comprehensive or accurate account of morality, and that the deontological properties of his own theory, though necessary, naturally point towards a dependence on 'some other method', which Sidgwick theorises to be utilitarianism. These conclusions set up the next chapter, which examines Sidgwick's account of that utilitarianism, and argues similarly to this chapter that whereas Sidgwick also attributes great value to the utilitarian principle, this too is incomplete without support from deontological concepts. Together, chapters 2 and 3 lay the groundwork for the thesis' central argument, developed in chapter 4, that deontological and utilitarian principles are necessarily combined in Sidgwick's theory of Rational Benevolence.

2.1. Sidgwick's Deontology in I:I : Initial Connections with the Intuitional View

Given *ME*'s infamous reputation as being somewhat difficult to read, its introduction is actually remarkably concise – and bold - in its approach to what are some inarguably substantial statements. It is in the introduction that Sidgwick draws attention to the fundamental problem in ethics – so similar to that outlined at the start of this thesis – that people generally aim to do things that they deem to be reasonable, and yet an enormous diversity of actions results from these these practical 'reasonings' (I:I:5-6, 14). It is here that he states that, for that reason, most single answers to the question 'why should I do something' only leave space for the same question to be asked from another angle (I:I:6). It is in the introduction that Sidgwick asserts, even at this early stage, that the term 'ought' needs to be handled carefully, given that it is most often used to describe actions that are relative to an optional end, when this is not the 'exhaustive' account of that concept (I:I:7). Further to this last point, it is in the introduction that Sidgwick argues that the pursuit of Happiness can be interpreted to be an ultimate end, in such a way that it is still prescribed by an 'ought' derived from Reason, and thus still accords with the Kantian categorical imperative (I:I:7); that is, it is in the introduction that Sidgwick states his own view of utilitarianism to be a version in which Happiness can be found to be an end *categorically prescribed* (I:I:8), at which people ought to aim.

These are no small claims. But from them, some minutiae can be derived that are vital for a full understanding of Sidgwick's interpretation of deontology, and of what he calls the Intuitional method. These are examined here in turn.

2.1.a. 'Good' and 'Duty' in Right Conduct

In the second section of I:I, Sidgwick draws a clear contrast between the general conceptions that Ethics is sometimes considered as an investigation into the true 'moral laws or rational precepts of Conduct; [and] sometimes an inquiry into the nature of the Ultimate End of reasonable human action – the Good...of man' (I:I:2-3). This is Sidgwick's distinguishing between the Aristotelian scholastic approach to ethics, which focused on the identity of the true 'good' for man and included eudaimonistic theories on how one was to live in order to attain that good, and the more modern approach to ethics, which specifically enquires into the nature of right

and wrong actions. Within the latter, it can be assumed that Sidgwick is including both deontological and utilitarian theories (i.e. the theories of Kant, Bentham, and Mill) – there is nothing in the text to state that by ‘right conduct’ Sidgwick means either one or the other. But an interesting insight arises here, even at this early stage, into Sidgwick’s view on ‘right’ action, and the role of good³⁶⁵. Good, Sidgwick says, can be assumed to be part of most discussions of ethics; ‘determination of precepts or directive rules of conduct’ are still necessary, if we require to know is *right* to actually *do*. Although Sidgwick is pointing out the historical decline of focus on the good here, there is a sense that Sidgwick is also demoting the concept in his own treatment of Ethics. He continues on to point out that the Intuitionist view – which he defines for the first time as that in which ‘conduct is held to be right when conformed to certain precepts or principles of Duty, intuitively known to be unconditionally binding’ (I:I:3) – gives very little weight to Ultimate Good at all.

Further, we commonly think that human good includes the performance of duty as well as the attainment of Happiness, and on this view again the performance of duty is not dependent on Ultimate Good. Not only then has the ‘good’, which Sidgwick loosely associates with Happiness here, been relegated in terms of its role in determining right conduct, but Sidgwick has also acknowledged that there is a place for a type of moral duty that is independent of both the good and Happiness. So whereas Sidgwick is not entirely specific at this point that this ‘duty’ belongs to the deontological position, his denial that the performance of duty can be conditional on knowledge of things achieved provides early evidence that Sidgwick understands ‘duty’ to be absolute, and unconditional³⁶⁶.

³⁶⁵ Generally, I will not capitalise ‘good’ in the thesis, apart from when it appears as part of Sidgwick’s phrase ‘Ultimate Good’. This is because Sidgwick is not always consistent about capitalising the word, and for me to capitalise it seemed to give it a certain significant weight that I do not necessarily intend. This is especially compounded by the fact that Sidgwick does not often capitalise ‘right’. *Ultimate* Good however, is representative of Sidgwick’s ultimate end of Happiness, and as Sidgwick *does* capitalise Happiness, this will appear more consistent. To the same end, I will also capitalise Happiness. The self-evident principles of Justice, Prudence, and Universal Benevolence (or just ‘Benevolence’, where appropriate) will be also capitalised, as per Sidgwick’s example, and in order to represent their particular status as Sidgwick’s self-evident philosophical intuitions – and the various forms of ‘Intuitionism’ are capitalised where it is clear that they are representing a certain type of epistemology, as opposed to Sidgwick’s *own* epistemic device (for example, within the context of Dogmatic Intuitionism). ‘Reason’, as representative of the particular cognitive moral faculty, is also capitalised, again in accordance with Sidgwick. Other than this, there are no other capitalisations.

³⁶⁶ This is as opposed to what could be a utilitarian understanding of duty, as I outlined in the introduction.

Sidgwick's emphasis on 'right conduct' (which here does not need defining as either the deontological or the utilitarian position), characterises the whole of *ME*³⁶⁷. And his preoccupation with what it is *right* to do is precisely what will lead him to the fundamental notion of moral obligation as being based in Reason, that I will argue causes him to align with a traditionally deontological, non-consequentialist understanding of 'duty'.

2.1.b. Further Outlines of 'Intuitionism' in I:I

Shortly after his discussion of the scholastic good, Sidgwick refers to the Intuitional view that he has just described as being the prevalent view. Sidgwick's reasons for

³⁶⁷ My reading of this passage as Sidgwick shifting the focus of ethics away from the good and towards the right is supported by *ME* when it is read as a whole. When this wider view of *ME* is taken, it is noticeable that Sidgwick is not greatly concerned with 'good'; there are only two chapters on the concept - I:IX, and III:XIV – whereas he is frequently occupied with the language of what is *right*. Having said this, it is important to have an overview of what Sidgwick considers good to be, as the concept is relevant to the Universal Good that is established at the core of Rational Benevolence. In I:IX, Sidgwick admits that 'virtuous' action may be determined by a moral ideal that is 'attractive rather than imperative', when something is judged to be 'good' in itself (I:IX:105). This is not, Sidgwick argues, a mere verbal substitution for 'rightness' – it is a fundamentally different form of intuitive judgement. This distinction serves to further delineate between Sidgwick's ideas of right and good: Right conduct involves an 'authoritative prescription' to do something, but it is not clear from the judgement that something is good that we ought to prefer that good to all other goods. Interestingly however, this leads Sidgwick to attribute to Ultimate Goodness a quality similar to that which he will attribute to rightness - what is ultimately good is *not* that which is simply good as the means to the attainment of some ulterior end. Through this, Sidgwick concludes that 'ultimate good' on the whole...must be taken to mean what as a rational being I should desire and seek to realise, assuming myself to have an equal concern for *all* existence' (I:IX:111-112). This 'good', which is good in itself, is independent of consequences (I:IX:113) – and this again accords with Sidgwick's later definition of rightness as *not* simply being what is fit to an end. This adds further support to my argument made below that Sidgwick has a predominantly non-consequentialist view of both certain moral *actions*, and certain *ends*, which is essential to my overall argument for the dual reasonableness of deontological and utilitarian values in Sidgwick's moral theory. It is also this temporally neutral and non-consequentialist notion of 'good' that is the basis of Sidgwick's self-evident principle of Prudence, and it is Prudence that establishes *Universal* Good as the rational end of moral action, within Rational Benevolence. Sidgwick's notion of Ultimate Good (or 'Happiness', as he more precisely defines it (I:IX:115, see also III:XIII:388,) may not receive quite the same level of attention as that of what is ultimately right, but it is a crucial component in his moral theory, that allows that theory to be seen as one of both the right *and* the good. In III:XIV (entitled 'Ultimate Good') Sidgwick considers whether, following his conclusions on Justice, Prudence, and Benevolence drawn in III:XIII (which I discuss at length below), the 'practical determination of Right Conduct' has actually come to rest on the determination of Ultimate Good after all (III:XIV:391). In this chapter Sidgwick concludes more robustly that Universal Happiness can be the only 'coherent account' of Ultimate Good, and that this does seem to indicate utilitarianism (III:XIV:406-407). This would seem to promote the role of the good *above* that of the right, as might be expected of a 'classical utilitarian' such as Sidgwick. But I argue here that these last comments from Sidgwick on the identity and role of good need to be read in light of those investigations in Book III, where it is seen that the Universal Happiness Sidgwick is referring to can *only* be established on a very precise basis, that depends absolutely on a notion of a rational and therefore categorical and unconditional *right*. I make this argument below, and in chapter 3 and 4.

believing this will become clear as he expands on this method of ethics, but for now we can look to Sidgwick's further definitions of the Intuitionist position. He describes consequences as not relevant to the 'rightness' of certain acts ('common moral opinion recognises...fundamental rules, e.g. those of Justice, Good Faith, and Veracity....as to binding without qualification and without regard to ulterior consequences' (I:I:8)). He also points to the role of Reason in the method. In this 'ordinary form of the Intuitionist view of Ethics', Sidgwick states, rules such as Justice etc. are prescribed as 'categorical', as the 'result of philosophical reflection'. It can be inferred from Sidgwick's linking of the term 'categorically' (where he attributes the term to Kant) to his phrase 'prescribed by reason' on the previous page he believes these rules to themselves be prescribed by Reason³⁶⁸. Sidgwick also argues here that to be prescribed by Reason 'categorically' means 'without tacit assumption of a still ulterior end' (a point that will shortly be of tremendous importance for understanding Sidgwick's own notion of moral obligation). This Reason-centric understanding of 'categorically prescribed' and the latent idea that right action is independent of consequences, are representative of Sidgwick's overall position on the deontological view³⁶⁹.

2.2. Reason, 'Ought', and the Deontological Version of Duty in Sidgwick

In *ME*'s introduction then, we have all the building blocks of what Sidgwick will develop throughout *ME* into his full understanding of the Intuitionist position. From this and other later definitions of that view, it is clear that Intuitionism is, for Sidgwick, the position that holds what is moral to take the form of duties/certain kinds of

³⁶⁸ Sidgwick himself makes an important clarification about this distinction at I:VI:83, where he acknowledges that 'almost any method may be connected with almost any ultimate reason...In my treatment of the subject, difference of Method is taken as the paramount consideration'. This adds weight to the argument that when Sidgwick capitalises 'Intuitionism', he is doing so in order that the word might specifically represent a certain method, which is the non-consequentialist duty-based one.

³⁶⁹ See also Sidgwick's table of contents at xxiii for another reference to Rules as prescribed without reference to ulterior consequences. Sidgwick does not actually use the language of 'Rules' very often in *ME*, but its appearance in the table of contents is helpful for further clarification that by this term Sidgwick is specifically referring to the non-utilitarian (i.e. non consequentialist) position. These definitions can be thought of as supporting Thomas Hurka's argument that Sidgwick has an 'extreme' view of deontology, by which he meant that Sidgwick had failed to recognise the possibility of prima facie duties, rather than only absolute ones (see this thesis chapter 1, pp.68-70). But it is my view that Sidgwick's interpretation of what appears to be a mainly Kantian form of deontology is not needlessly extreme, but rather simply the most accurate description of the form of deontology that was prevalent in his time.

actions which are prescribed unconditionally, and without reference to ulterior results or consequences³⁷⁰. I believe then that it is no great step to equate this non-consequentialist emphasis with the deontological view of ethics, as it was outlined in the thesis introduction. However, what is significant about my claim is not necessarily that Sidgwick's 'Intuitionism' represents the traditional deontological approach to morality, but that this type of duty that characterises this Intuitionism (with a capital 'I') is also the *same* duty that informs Sidgwick's own theory of moral obligation, that will later appear in his epistemic intuitionism (with a small 'i'). As we have reached Sidgwick's discussion of rightness in I:III, the first part of the argument for this claim can now be given. I aim here to show that Sidgwick's particular use of Reason, and his subsequent understandings of 'ought' and 'right', create a recognisably deontological theory of moral obligation whereby moral action is categorically and unconditionally prescribed (mainly) without reference to consequences - and that this notion of duty remains categorical and unconditional and thus *deontological*, even when it appears in Sidgwick's more utilitarian outline of rightness³⁷¹.

³⁷⁰ Sidgwick gives further definitions to this effect at I:III:20, 26, I:VIII:96, 98, I:IX:105, and III:I:200 – the significance of these are discussed below. Sidgwick's own use of intuitionism as his method of epistemology may occasionally be blurred (for reasons I am about to suggest), but his definition of Intuitionism with the capital 'I' is clearer than perhaps David Phillips allows (Sidgwick is also clear about his own opinion on intuitionism (with a small 'i'), as opposed to its representing just one ethical system, elsewhere in ME. I:VIII is specifically dedicated to discussing the various forms of intuitionism, and III:I:201 shows Sidgwick to be aware of the possible intuitive basis of all three methods.) Furthermore, in his response to Calderwood, Sidgwick does reiterate his 'fundamental assumption of Intuitionism' – as he intends it to be interpreted in *ME* – which is 'that we have the power of seeing clearly what actions are right and reasonable in themselves, apart from their consequences' ('Professor Calderwood on Intuitionism in Morals', in, *Mind* Vol.1, No.4 (1876), p. 565). This fully supports my argument that by 'Intuitionism', Sidgwick is referring to the non-consequentialist, deontological method. Whereas Sidgwick states in the Calderwood paper that in his own opinion intuitionism does inform at least both utilitarian and 'Intuitional' (deontological) methods, even the passage from which Sidgwick has restated that definition to Calderwood (III:I:200) also makes reference to the concept of unconditional duty.

³⁷¹ I do need to draw attention here to an alternative view on Sidgwick's Intuitionism (with a capital 'I'). Alan Donagan ('Sidgwick and Whelwellian Intuitionism; Some Enigmas' (*Canadian Journal of Philosophy*, Vol. 7, No.3 (1977), 447-465) states that at I:VIII:101, Sidgwick gives what appears to be a Whelwellian version of Intuitionism (the version that I represent with a capital 'I'). This is also the Intuitionism that Sidgwick associates with Dogmatic Intuitionism, or common sense morality ("It is such a system as this which seems to be generally intended when Intuitive or *a priori* morality is mentioned, and which will chiefly occupy us in Book iii" I:VIII:101). Within this Intuitionism, 'actions are unconditionally prescribed without regard to ulterior consequences' (I:VIII:98). If it is the case that it is Whelwell's Intuitionism that Sidgwick has in mind at all times when he describes this moral-theoretic (if I may borrow that term from Phillips - see *Sidgwickian Ethics*, p.6) and *deontological* position, then this would seem to conflict with my claims that there is a Kantian element in that Intuitionism, and that Sidgwick retains this Kantian aspect of the Intuitional (deontological) method in his own philosophical intuitions that constitute his moral theory. My response to this is in three parts. The first is to suggest, contra Donagan, that Sidgwick did have Kant's position in mind as a partial informant of the Intuitional view. This seems a reasonable argument to make, as we see quite clearly earlier in *ME* Sidgwick

2.2.a. Reason, Right, and 'Ought': Sidgwick's Appeal to Kant

In *ME* Book I, Chapter III ('Ethical Judgements') Sidgwick returns to his points made in the introduction regarding 'actions that we judge to be right', the idea that what 'ought' to be done is 'reasonable', or 'rational', and of ultimate ends as being prescribed by Reason (I:III:23. See also I:IX:105). How far, Sidgwick asks at the start of this chapter, is Reason responsible for our judgements of what is 'right', and of what we 'ought' to do? Reason is often recognised to be a part of the moral process, but serious contention arises (at Sidgwick's time) from Hume's claim that 'Reason....can never of itself be any motive to the Will'³⁷². Sidgwick, however, follows Kant's version of the importance of Reason in morality. It is from this defence of Reason and rationality as integral to moral judgements that Sidgwick's own interpretation of moral obligation begins to emerge.

There are substantial grounds in I:III on which to argue that there is a Kantian influence on Sidgwick's view of Reason. Contrary to Hume, Sidgwick uses Reason

outline a specifically non-consequentialist duty-based form of Intuitionism, in which duties are prescribed by Reason. Secondly, I agree with Donagan that Sidgwick *is* rejecting the rawer Whelwellian version of epistemic intuitionism throughout Book III (it is intentional that I have used the lower case 'intuitionism' here – the passage on 101 is a clear example of the confusion that can sometimes arise when Sidgwick's various uses of moral-theoretic and epistemic intuitionism converge). The PD clearly shows Sidgwick rejecting *Whelwell's* version of Intuitionism (there is again convergence between the two types of intuitionism here), on the basis that its principles were 'hopelessly loose' (xv). So what Sidgwick is rejecting throughout (most of) Book III is the non-self-evident, 'hopelessly loose' principles of Dogmatic Intuitionism, or common sense morality. However, the final part of my response to the Whelwellian/Kantian 'Intuitionism' issue is to refer *back* to my first point, that Kant was a partial informant of that non-consequentialist position, and argue that *this* particular element of moral-theoretic Intuitionism *does* appear in Sidgwick's own epistemic intuitionism. That is, whereas it may be Whelwell's semi-moral-theoretic/semi-epistemological intuitionism that Sidgwick is rejecting in Book III, Sidgwick adds the Kantian element (which retains non-consequentialism) back in when he argues for his own philosophical intuitions, and this helps make sense of Whelwell's position/Dogmatic Intuitionism. This suggestion is directly supported by Sidgwick's account in the PD. There, we see him immediately reject Whelwell's Intuitionism (see also Sidgwick's rejection of Whelwell in Sidgwick's *Outlines of the History of Ethics* (London: Macmillan: 1886) pp. 233-4). We then see Sidgwick 'add' a Kantian intuition to his moral theory, despite his 'early aversion to Intuitionist Ethics derived from the study of Whelwell' (xvi). Sidgwick had, in his words, 'become an Intuitionist to a certain extent' (xix). *This* - the addition of a Kantian intuition (which *also* has the non-consequentialist quality) - is what allows the reconciliation between Whelwellian Intuitionism (or Dogmatic Intuitionism) and utilitarianism. Thus Sidgwick *improves* on Whelwellian intuitionism, with Kant, and with his own philosophical intuitions. (Incidentally, Donagan agrees that Sidgwick would have wanted to *reject* working out this 'improved dogmatic intuitionist system on Whelwell's lines' ('Sidgwick and Whelwellian Intuitionism' p.460).

³⁷² What is at stake in opposition to Reason in Hume's view are the 'irrational' desires that arise from bodily functions, such as hunger, or emotions, such as anger. Sidgwick disagrees with Hume on the basis that the notions represented by the words 'ought' and 'right' are entirely different from notions derived from physical experience of the world (I:III:25). Sidgwick then moves on in I:III from this starting point to develop this idea more fully.

to ‘denote the faculty of moral cognition’ (I:III:34), which he represents with the maxim ‘what I judge ought to be must, unless I am in error, be similarly judged by all rational beings who judge truly of the matter’ (I:III:33). This is (as will be confirmed later by the PD), obviously resonant of Kant’s fundamental maxim ‘act from a principle or maxim that you can will to be a universal law’. In Sidgwick’s view on other concepts too – his statement that the meaning of ‘right’ and ‘what ought to be’ cannot be ‘fitness to some ulterior end’ (I:III:26), his argument that the concepts of ‘ought’ and ‘duty’ that make up moral obligation have a peculiar nature in the context of ethical judgements, and his belief that ‘right’ and ‘ought’ are ‘too elementary to admit of any formal definition’ (I:III:32) – we see developing a position that does not seem to accord entirely with the utilitarian view that morality lies in consequences. This particular understanding of ‘right’ as *not* meaning fit to an ulterior end is a direct result of Sidgwick’s understanding of the role of Reason in the establishment of what is ultimately *right* being that which is ultimately *rational*, and needs no reference to anything further than itself³⁷³.

Even more importantly for how this understanding of rightness and rationality translates into moral obligation, Sidgwick shares Kant’s views on the relationship between Reason and volition. Sidgwick states that a cognition that something *ought*

³⁷³ The place that Sidgwick affords to Reason is, at this stage of *ME*, slightly weaker than Kant’s. Where Kant propounds Reason as the *only* possible source of and for morality, Sidgwick says at I:III:33 that he does not ‘mean to prejudge the question whether valid moral judgements are normally attained by a process of reasoning from universal principles or axioms...it is not uncommonly held that the moral faculty deals primarily with individual cases as they arise, applying directly to each case the general notion of duty, and deciding intuitively what ought to be done’. But Sidgwick continues on to explain that as this points to something more like ‘Sense’, rather than cognition, there could be no way of telling whether A or B is right in any given situation, and Sidgwick is ultimately searching for what can finally be called morally valid. From this we can assume that Sidgwick is searching for some kind of *objective* view of what is moral – and for this Reason is the most appropriate route to follow. (He reiterates this point in a footnote at I:VI:77, stating again that not all people will determine what is right conduct on the basis of what they consider to be *reasonable* – but that he is restricting his investigation in *ME* into what *is* prescribed by Reason, and that he intends to ‘exhibit the self-evident practical axioms’ that he considers to be included in that category). Sidgwick has also just outlined a loose version of Kant’s maxim, which reads ‘what I judge ought to be must, in unless I am in error, be similarly judged by all rational beings who judge truly of the matter’ (I:III:33), stating that this is what the moral faculty is ultimately pointing to, when we consider the notion of ‘ought’. Despite Sidgwick’s less obviously rigorous adherence to Reason then, it is clear that Sidgwick believes Reason to take the same form as in Kant’s doctrine – and to yield the same results in terms of what is meant by ‘ought’. Whereas Sidgwick’s investigations into the reasoned convictions of ordinary people reveals the fact that they are actually not ultimately self-evident, and therefore not ultimately reasonable in themselves, Sidgwick eventually pursues what *can* be called ultimately reasonable in order to arrive at his version of the Kantian maxim, and at the principle of aiming at the general happiness. See I:VIII:100-101, III:XI:337-345, III:XIII:379-389 – and my chapters 3 and 4 for the full development of this point.

to be done naturally brings with it the impulse to action (I:III:34, 37)³⁷⁴. He also agrees with Kant that it is only rational for a rational being to obey dictates that they perceive to be the result of Reason (I:III:34)³⁷⁵. From these two precepts it can be directly inferred that Sidgwick's view of Reason leads him, as Kant's view did for him, to the idea that actions prescribed by moral 'oughts' are unconditional³⁷⁶. This Reason-based idea of unconditional moral obligation then pervades I:III. Once Sidgwick has drawn his conclusions on the nature of 'ought', the language of 'dictates of Reason' appears frequently – (at I:III:34 twice ; at I:III:36, also twice). From then on, this language resolves itself into the more precise vocabulary of unconditionality, unconditional imperatives, 'unconditionally prescribed by Reason', and – in a final strong indication of the correlation between Kant and Sidgwick on this point – the categorical imperative (I:III:35, 37). The 'ought' - which always represents duty in some form - on Sidgwick's view thus takes on a distinctly non-consequentialist and *deontological* appearance.

2.2.b. The Deontological Basis of Sidgwick's Theory of Moral Obligation

We can now return to the definition of 'deontological' given in the thesis introduction and re-iterated above (as the view of morality that holds moral value to lie in the performance of certain duties that are right unconditionally and irrespective of consequences), in order to ascertain how far Sidgwick's concept of 'ought' and 'duty' accords with it. As just noted, both Sidgwick's language and his concepts are, in these pages, remarkably deontological according to the *first* part of the definition, which stipulates unconditionality, and categorical imperatives. But a question arises

³⁷⁴ Sidgwick also argues that 'ought' must mean that one is able to perform a certain action – it is not possible to conceive that one ought to do something from which one is prevented from actually doing (I:III:33). This view fully supports John Deigh's internalist reading of Sidgwick, and argues against David Brink's externalist account (see this thesis chapter 1, p.161).

³⁷⁵ Actions that are 'reasonable' in this way, Sidgwick equates with actions that are 'right' 'in an absolute sense' (I:III:35).

³⁷⁶ Sidgwick does not specifically state in this chapter that he believes unconditional actions to apply universally. On this point, Sidgwick's position differs to Kant's, for whom universalizability is a crucial component of judgements made within the categorical imperative. Also, Sidgwick indicates at other times (IV:IV:466-468, 473-474) that he is a cultural relativist, which also challenges the notion of universalizability. However, it is not difficult to infer from Sidgwick's emphasis on the unconditionality of oughts that he has at least a notion that these imperatives are universal. Furthermore, Sidgwick does return to universalizability in his own moral theory – and he does so *specifically*, as seen in the phrase 'Universal Benevolence'. This is the direct result of his following Kant's argument that ultimate rationality establishes categorical oughts. That is, Sidgwick will eventually derive a universalising stipulation from his non-consequentialist version of moral obligation.

regarding the traditional deontological position's rejection of consequences. Sidgwick's language of 'unconditional' and 'categorical' *implies* a move away from the consequentialism inherent in utilitarianism, but this hardly counts as solid evidence. There are actually two points that the issue of consequentialism raises against my claim just made about Sidgwick's notion of 'ought', which render that claim as yet incomplete. The first is that deontological 'duty' (as opposed to utilitarian understandings of duty) is *specifically* defined as non-consequentialist. The second is that Sidgwick has stated at I:III:26 that this notion of 'ought' which I am arguing represents a (primarily) deontological understanding of duty also applies to 'the adoption of certain ends'. The idea of ends obviously involves the idea of *means* to those ends, which in turn appears to bring moral value to rest on whether or not those ends are achieved - and this is a distinctly consequentialist, i.e. non-deontological, position. At the very least, Sidgwick seems to be arguing that his understanding of duty can be utilised as much by the utilitarian, consequentialist view. After all, the view that Reason will always generate moral obligation – or a moral duty – does not need to be exclusively deontological. How, then, is this appearance of the pursuit of rational *ends* to be made congruent with an argument that Sidgwick's is a traditionally deontological version of duty, that excludes consequences?

I believe that it is more than possible to argue that Sidgwick's version of moral obligation *is* a deontological one, or at least so close to it that it would be more appropriate to understand it as deontology than utilitarianism. The argument depends firstly on understanding ultimate, rational ends to themselves be non-consequentialist in nature, and secondly on Sidgwick's particular use of Reason (i.e. the Kantian version) and the unconditionality/categorical imperative on action that this produces. Both of these points are most conveniently explained by briefly revisiting the steps through which Sidgwick defined 'right' and 'ought'.

At I:III:26, Sidgwick immediately dismisses 'right as fit to an ulterior end'. He points out that we recognise some actions to be right unconditionally, that we '*similarly* regard as 'right' the adoption of some ends' (italics mine), and that on this basis the idea that 'rightness is an attribute of means, not of ends' is 'clearly inadmissible'. On the basis of Sidgwick's rejection of rightness as meaning fit to an end, his rational ultimate *ends* themselves thus take on a non-consequentialist character. That is, as they are ultimately rational, they are also ultimately right, in and

of themselves, with no reference to anything further, such as ulterior consequences. It is only in this way that an *end* can be truly called right, given Sidgwick's definition of rightness - and as can be seen from his two stipulations of rightness on page 26 both certain actions and certain ends clearly share the non-consequentialist factor of being right 'unconditionally, without regard to ulterior results'.

From here, Sidgwick goes on to make his well-known argument that as certain actions *and* certain ends are right in this particular, ultimately rational way, the notion of 'ought' must apply to both, in the form of a categorical imperative (I:III:35-36). Sidgwick admits that one might argue there is no obligation without reference to consequences, but, he argues, these consequences pertain to ultimately reasonable (and universal) ends, and the recognition of an ultimately rational end incurs the same categorical 'ought' as that which applies to certain kinds of actions, and which has its basis in Reason³⁷⁷. Means (to those ultimate ends) are included to some extent, but those means are only established on the original non-consequentialist basis of the rational end³⁷⁸. Permeating both the end and its action then, there is a categorical and ultimately non-consequentialist understanding of rightness. It is this that I argue informs Sidgwick's whole view of moral obligation, and moral duty.

³⁷⁷ '...they recognise some universal end', Sidgwick says, '...at which it is ultimately reasonable to aim. But in this view, as I have before said, the unconditional imperative plainly comes in as regards the end, which is – explicitly or implicitly – recognised as an end at which all men 'ought' to aim; and it can hardly be denied that the recognition of an end as ultimately reasonable involves the recognition of an obligation to do such acts as most conduce to the end' (I:III:35). He concludes this point on the following page: '...the notion of 'ought' as used in either dictate is that which I have been trying to make clear' (I:III:36).

Sidgwick does point out here that the obligation is not unconditional in quite the same way, but that it still does 'not depend on the existence of any non-rational desires or aversions'.

³⁷⁸ Kant himself also provides some remarkable agreement with Sidgwick on this very point about certain ends. In the *Critique of Pure Reason*, Kant discusses the role and identity, in his view, of the good. The 'highest good', he argues, is indeed happiness - but *only* Happiness as conceived by rational agents (*Critique* (1781) 6:5-7). We have a moral duty to pursue this highest good, but for Kant this is not an additional duty, but the sum of all other duties (6:5). The parallel between this and my argument that as rightness can only lie in what is ultimately rational and not merely fit to an ulterior end duty can only truly be assigned to the ultimate rational *end* of Happiness and not its means is clear. It is also clear that Sidgwick does not attribute real moral value to *means* to ends. Means in themselves are not ultimately rational, therefore they cannot be unconditionally and categorically right in the way that their ends are. i.e. they cannot be *duties*. We might also add that Sidgwick never once refers to 'consequences' here.

2.2.c. Consequentialism in Sidgwick's Theory of Moral Obligation

In the previous section, I have been arguing that Sidgwick's view of rightness is non-consequentialist in that he does not define as right that which is simply fit to an end. Further, as rightness is attached to what is ultimately rational, 'ought' becomes a categorical and unconditional imperative. This might, however, be thought to not fully shift the onus of moral value away from consequences. I do admit that Sidgwick's inclusion of *ends* allows a degree of utilitarianism into this definition of duty, and it must be recognised that Sidgwick does at one point attach the 'ought' to the idea of a 'fittest means' to an end (I:III:36). But to this objection I have two responses. The first is that even when Sidgwick admits that his notion of ought/moral obligation applies to the utilitarian position as well, which he does when he says that nothing he has said is intended in favour of the Intuitional view as opposed to the deontological one (I:III:35), it still has with it the very particular interpretations of 'ought', and 'rightness', which can only come about on the basis of recognising an ultimately rational end, which as such pertains to nothing further and so is non-consequentialist in this way³⁷⁹. That is, even the way in which utilitarian dictates share in Sidgwick's notion of moral obligation depends on the prior establishment of that obligation as unconditional, categorical, and not 'right' merely on the basis of consequences³⁸⁰. And that, I argue, accords with the unconditional form of duty that is central to the traditional deontological understanding of morality. Here then, we have a crucial prediction of how Sidgwick's understanding of duty and rightness retains a

³⁷⁹ The timing of this passage is actually still supportive of the argument that there is a distinctly deontological emphasis in this chapter of *ME*. It appears after Sidgwick makes his point that unconditional imperatives are still relevant to the pursuit of ends – the impression is that Sidgwick himself has recognised that he appears to have taken a predominantly deontological line in this chapter.

I should also point out here that I do not mean to argue that an end such as Happiness, or Universal Good, is non-consequentialist in its *dictates*. I am arguing that the *end itself*, being rational, is an ultimate one and as such is 'held to be right unconditionally, without regard to ulterior results' (I:III:26).

³⁸⁰ This argument is further (and compellingly) supported by Sidgwick's briefer account of how Reason applies to ends given in *ME*'s introduction, when he specifically attaches 'ought' to the idea of the categorical imperative. In light of Sidgwick's argument for Reason/unconditional moral obligation in I:III, we can now see more clearly the importance of Sidgwick's statement at I:I: 8 about Happiness as an end 'categorically prescribed'. On these pages, we saw Sidgwick mark a clear distinction between the position that uses 'categorically', and the utilitarian position. But then he says that if aiming at the general Happiness is a duty, then the end of Happiness becomes an ultimate end categorically prescribed, of exactly the same kind as he had just discussed in relation to the Intuitional view. In this way, 'ends' and 'certain kinds of action' share exactly the same quality of categorical obligation.

deontological character when Sidgwick uses it in his own moral theory, even when that theory yields a utilitarian principle³⁸¹.

The latter point leads to my second response to the objection about consequences, which is that the minor role that consequences might be argued to play in Sidgwick's theory of moral obligation is actually not detrimental to my argument for Sidgwick's reconciliation between deontological and utilitarian principles in Rational Benevolence. Perhaps we cannot say absolutely that Sidgwick's notion of duty excludes consequences completely³⁸². But as Sidgwick says, no method can – not even the Intuitional one (I:VIII:96-97)³⁸³. It is integral to the specific position I argue Sidgwick to occupy between deontology and utilitarianism that 'ought', as established by Reason, *is* in fact equally relevant to systems of morality that posit ends as motives for moral action as to systems that propound non-consequentialist duties. In *both* cases, the 'ought' is unconditional – because in both cases, the ought is established by what is ultimately reasonable. Again then, there is a clear prophecy here of the ability of Sidgwick's notion of moral obligation to transcend the traditional deontological/utilitarian boundary, on the basis of the *shared appeal to Reason* of both positions, and this is what I argued in the thesis introduction is needed, to account for the dual reasonableness that causes moral ambivalence. But it is my firm view that it is the non-consequentialist, deontological notion of rightness and duty (understood as either actions or ends that are unconditionally and categorically prescribed *without* reference to anything further) that is predominantly informing Sidgwick's theory of obligation in this most important of chapters. The rightness of those ends is based on the rational properties intrinsic to that end – not on anything that they bring about. The rational 'ought' is always, on Sidgwick's view, a form of unconditional and categorical duty – and this understanding of duty applies just as much to ends which are right *without*

³⁸¹ It is of course that argument that is crucial to the establishment of Rational Benevolence as a synthesis, in which deontological and utilitarian properties retain their own independence, while also relying on each other for full validation.

³⁸² See chapter 3 of this thesis for an expanded account of Sidgwick's reintroduction of 'means' to ends.

³⁸³ The denial of consequences cannot be extended to the whole range of duty, Sidgwick argues on p.96, because 'no morality ever existed which did not consider ulterior consequences to some extent'. In order to differentiate between specifically non-consequentialist and specifically consequentialist positions, Sidgwick states that 'we must understand that the disregard of consequences, which the Intuitional view is taken here to imply, only relates to certain determinate classes of action (such as Truth-speaking)'. But it is clear that Sidgwick does not believe consequences to be completely irrelevant to the Intuitional view, and this further reduces the gap between the two positions.

reference to ulterior consequences, as it does to the commonly held duties such as Veracity, and Justice.

2.3. Sidgwick's Expansion on Deontology as it is Commonly Held

It has now been seen that Sidgwick believes moral rightness, and the moral 'ought', to lie in what is unconditional and categorically prescribed by Reason. I have also argued that as rightness does *not* lie in fitness to an ulterior end, this is a mainly non-consequentialist and deontological interpretation of moral obligation. Thus we already have a strong impression that a certain deontological understanding of moral duty is, in some way, important to Sidgwick. As for where this deontological understanding of moral obligation features in Sidgwick's own wider moral theory, that discussion is temporarily delayed in accordance with the order of Sidgwick's points in *ME*. For ahead of Sidgwick's argument for his own theory of Rational Benevolence is his investigation into the role that deontological principles have in the moral processes of ordinary people, and his conclusions on this are essential foundations to that theory. In these sections then, I examine Sidgwick's account of common sense morality – or 'Dogmatic Intuitionism' – with the purpose of drawing out both the value that he ascribes to this method of ethics, and the ways in which he believes it to be insufficient. This is done in preparation for the third and final stage of the chapter, which examines Sidgwick's incorporation of deontological principles within his own work.

2.3.a. The Place of Common Sense Morality

Here, some further observations on Sidgwick's definitions of Intuitionism are required, as these will be important to the final stage of my argument in this chapter, that elements of the Intuitionist view appear in Sidgwick's epistemic intuitionism. They also serve to create a wider picture of the place of common sense morality in Sidgwick's thought.

In I:VI, Sidgwick provides a convenient summary of his points on ethical systems made so far, reiterating that ultimately valid reasons for acting tend to be supplied either by a belief in certain notions (such as Happiness) as ultimate rational ends, or by an adherence to 'Duty, as prescribed by unconditional rules' (I:VI:78). It is this

latter position that Sidgwick defines as Intuitionism (understood as the system of ethics, rather than Sidgwick's epistemic version) and then associates with 'common sense morality', in Intuitionism's 'Dogmatic' form, at I:VIII:101³⁸⁴. In doing so, Sidgwick also removes any ambiguity from how 'duty' is now to be interpreted³⁸⁵. Sidgwick furthers the association between common sense morality and deontologically-based Intuitionism throughout I:VIII, explicitly stating that he has "used the term 'Intuitional' to denote the view of ethics which regards as the practically ultimate end of moral actions their conformity to certain rules of dictates of Duty unconditionally prescribed" (I:VIII:96), and elsewhere adding 'without regard to ulterior consequences' to a very similar definition (I:VIII:98)³⁸⁶.

³⁸⁴Sidgwick perceives three distinct forms of the Intuitional method, through which its epistemology imparts its edicts. These are 'Ultra-Intuitionism' (or 'Perceptual Intuitionism'), 'Dogmatic Intuitionism', and 'philosophical intuitionism' (I:VIII:99-102). Importantly, Sidgwick considers all three manifestations to be active in our moral processes, but Sidgwick associates common sense morality with the second form, Dogmatic Intuitionism, which holds that we can 'discern certain general rules with really clear and finally valid intuition' (I:VIII:101).

The concept of 'common sense morality' has appeared earlier in *ME* – in the first and second prefaces (v, vii), in I:I when Sidgwick talks about the 'practical reasoning of men generally' and 'the thought of ordinary men', and in its exact phrasing when Sidgwick outlines the question central to *ME*, 'what are the different practical principles which the common sense of mankind is *prima facie* prepared to accept as ultimate?' (I:I:6). In the passage on p.6, Sidgwick uses the term 'common sense morality' more broadly to refer to both absolute-duty based and end based methods of moral reasoning. In I:VI, Sidgwick specifically restricts the term so that it refers only to the former (I:VI:85, 87).

³⁸⁵ This firm distinction is something of a contrast to the earlier chapters, where – as was seen – Sidgwick's view of Reason infused both absolute-principle based and end based moral systems with a sense of 'duty', brought about by 'ought'. Sidgwick is at his clearest by this point in *ME* that he considers the moral position that supports *duties* to be that which is associated with non-consequentialist unconditionality (I:VI:77, 78, 81, 84, see also I:VIII:96, 98), but this appears to be a more formal, academic presentation of the position – made for the sake of clarity – rather than Sidgwick's own opinion. It would seem as though now Sidgwick is moving into the stage at which he intends to examine the methods as they appear in ordinary moral processes (as opposed to his examining them as parts of his own moral philosophy), he is drawing more distinct lines between the two positions. Sidgwick returns to considering absolute-principle based *and* end based types of morality to share this form of duty in his chapter 'Philosophical Intuitionism' (III:XIII), where he develops his own moral theory. This is of course the main focus of this thesis' chapter 4.

³⁸⁶ See also I:VIII:101-102, and III:I:215, where Sidgwick states the 'public opinion of the community', 'positive morality of the community', and 'the consensus of mankind' to be 'more significantly termed 'the morality of Common Sense'.

Sidgwick also gives a particularly distinct definition of Intuitionism at III:I:200, which effectively summarises his entire position in one passage: '...in such maxims as that duty should be performed 'advienne que pourra', that truth should be spoken without regard to consequences, that justice should be done 'though the sky should fall', it is implied that we have the power of seeing clearly that certain actions are right and reasonable in themselves, apart from their consequences...such a power is claimed for the human mind by most of the writers who have maintained the existence of moral intuitions; I have therefore thought myself justified in treating this claim as characteristic of the method which I distinguish as Intuitionism' (III:I:200). This passage has particular significance for showing that Sidgwick, as an intuitionist himself, must to some extent have to exclude consequences from his epistemology. Singer and Lazari-Radek make a similar point in relation to this same passage, that supports my argument (made in full below) that Sidgwick excludes consequences from his initial intuitions, at least to an extent. They emphasise that Sidgwick is specific in this passage that

It is interesting that Sidgwick should create this direct connection between ‘common sense morality’ and the Intuitionist (or deontological) view. The implication, which is emphasised by Sidgwick’s use of the language of ‘intuitionism’ and ‘intuitions’ to represent the position that upholds unconditional, absolute rules of duty, is that mankind’s morality is actually by nature (to some extent) deontological³⁸⁷. That is, in its rawest form, before it is passed through the rigorous testing and examination that Sidgwick intends in *ME*, our general, everyday practical moral reasoning – our immediate moral *intuitions* – originate from non-consequentialist, unconditionally binding principles³⁸⁸. The very term ‘common sense

intuitionists cannot refer to consequences (*Point of View of the Universe*, pp.69-70). If Sidgwick is going to call himself an intuitionist, then he too is prevented from using consequences in his epistemic method of acquiring of moral truths. Singer and Lazari-Radek’s point seems to be that if Sidgwick is an intuitionist in the wider (epistemic) sense, then he must, by default, to some extent also be an Intuitionist in the narrower sense. But they do point out that Sidgwick is not completely consistent about the exclusion of consequences (p.70), and that calculation of consequences must necessarily appear in both Prudence and Benevolence. It will be seen below (and in chapter 4) that I take a very different view of the consequentialism of those two principles; I do not dispute that they are consequentialist in their dictates, but I do dispute that their moral property, or *nature* is consequentialist, arguing instead that they are recognisably deontological in nature. This is what allows me to make a stronger argument than Singer and Lazari-Radek that Sidgwick does actually retain a significant degree of consequentialism-exclusion in his own intuitionism. Further to the passage at III:I:200, Sidgwick elsewhere states the Intuitionist view to refer to the results of certain ‘classes’ of actions, in which the nature of the act (as right) has already been determined. It is in the case of these actions, such as Truth-speaking, for example, that consequences can be disregarded from a moral judgement. This is part of Sidgwick’s delineation between Intuitionism and consequentialist based moral systems such as utilitarianism, that he makes in the wake of his point that there is no system of morality that does not consider ulterior consequences to some extent (I:VIII:96) – and it is important that he does make this clarification. Had he not done so, his statement that there is no morality that does not consider consequences to some extent could possibly be read as a clandestine assertion that all morality is found to be utilitarian in the end. This is not what Sidgwick means here, and his example that one might know that the truth one is about to speak to a jury will lead them to the wrong conclusion, and yet still actively believe speaking the truth to be right, regardless of that consequence (I:VIII:97), is evidence that Sidgwick sees a clear difference between ethical methods that hold actions to be right per se, and those that hold actions to be right by virtue of their results.

³⁸⁷ It is not even as if common sense morality could be thought of here as being a sort of hybrid collection of utilitarian and deontological principles, which might make it more understandable that Sidgwick would consider *this* position to best represent the general reality of ordinary morality; Sidgwick maintains a clear position that common sense and utilitarianism are very much distinct, and that utilitarianism is the alternative to common sense, not a part of it. See I:VI: 83-84, 85-87, and I:VIII:101-102).

³⁸⁸ Despite the general opinion that Sidgwick’s own intuitionistic epistemology is utilitarian, I:VIII unequivocally shows Sidgwick arguing that the epistemological *method* of intuition (that is, intuitionism with a small ‘i’ as I have suggested we understand the distinction) does also belong to Intuitionism with a capital I – i.e. to the deontological moral view that actions should conform to ‘certain rules or dictates of Duty, unconditionally prescribed’ (I:VIII:96). When justifying his particular use of intuition in this context (I:VIII:96) Sidgwick does grant that ‘intuitive knowledge’ cannot *absolutely* discount reference to consequences (see fn.24 above), and it is at this point that he goes on to state that as ‘pleasure as a reasonable ultimate end of human action’ (the belief integral to the Hedonistic view) can only be known intuitively and not inductively, then Hedonism can also be called an intuitionist method (I:VIII:97-98. See also p.102). But Sidgwick then clarifies that for the purposes of *ME* he is going to follow the ‘prevailing opinion of ordinary moral persons, and of most writers who have

morality' suggests even more strongly that this absolute-principle based moral position is the most familiar, the most conventional – and the most typical of ordinary people³⁸⁹. It should also be noted that thus far in *ME* there has been a certain dominance by the deontological view, and not the utilitarian. This is testimony to my argument that *ME* is not merely a treatise on utilitarianism, and that a significant proportion of *ME* is actually concerned with the opposite position. This is not to say positively that Sidgwick is defending the deontological view, or that he does not defend the utilitarian view. We have yet to confirm the value that Sidgwick does really ascribe to deontological principles – and Sidgwick certainly has not argued that common sense morality is a definitive or even entirely dependable system of ethics³⁹⁰. But what is being said, and it appears to be the pertinent time to do so given Sidgwick's long preoccupation with the Intuitionist view to which I am drawing attention in this chapter, is that the deontological view of morality is important to Sidgwick, in one form or another, throughout *ME*.

Continuing with the theme of Intuitionism's importance, Sidgwick asserts at I:VI:87 that he will address Intuitive Morality ahead of Utilitarianism on account of its being more immediately comprehensible. Having identified that Dogmatic Intuitionism, or the morality of common sense, is generally what is meant by the term 'Intuitive' morality, this is the method with which Sidgwick will be concerned in Book III. Specifically, this is because these rules require for true precision further contemplation and abstract moral consideration (I:VIII:100-101). This contemplation, Sidgwick believes, will go some way to removing the doubt and conflict that

maintained the existence of moral intuitions' (I:VIII:98) in defining Intuitionism as the belief that there are certain kinds of actions unconditionally prescribed without regard to ulterior consequences. And Sidgwick seems to have no trouble in this chapter with expanding widely on how epistemological intuitionism applies specifically to the Intuitionist, or deontological, method. Even in I:VI, Sidgwick uses 'intuitively' as the adverb attached to the enunciations of 'the Common Sense of mankind' (I:VI:85). It may not be absolutely certain that it is Sidgwick's own view that intuitionism with a small 'i' as an epistemic device applies *predominantly* to the deontological position – but it is certain that Sidgwick has aligned himself with this position in *ME*, and that he has called the deontological view after this very method. In any case – and more importantly for the later purposes of this thesis - there are grounds on which to argue that Sidgwick believes the method of intuitionism to lead to *both* the deontological and the utilitarian positions. It is my view that Sidgwick does hold epistemic intuitionism to be relevant to both the deontological and the utilitarian position, as he describes in the PD (this is seen below) and that this is a crucial component of the argument that Sidgwick eventually affirmed both deontological and utilitarian principles to be intuitive, and essential to morality. It is on this basis that I claim, in disagreement with David Phillips, that Sidgwick's intuitionism is not merely utilitarian.

³⁸⁹ Sidgwick also considers common sense morality to be the more accessible form of ethics, and the least demanding, or confusing, when compared to utilitarianism (I:VI:87), although these are not necessarily value judgements of the two positions, but rather observations.

³⁹⁰ See I:VI:82-83 for example, and I:VIII:100 -103.

inevitably (and obviously) occurs between individual moral judgements. For this reason, Dogmatic Intuitionism will actually eventually become the third type – Philosophical Intuitionism – which accepts the morality of common sense as its basis, but searches for a ‘undeniably true and evident’ philosophical basis for it which that common sense morality itself does not provide (I:VIII:102). This transition from dogmatic to philosophical Intuitionism is precisely the process that Sidgwick himself follows, in attempting to elucidate the precepts of the former, and concluding that the true self-evident basis for these precepts actually lies elsewhere³⁹¹. It is to that process, through which Sidgwick ascertains the value of Intuitionism, that this chapter now turns. But first, we can now conclude more fully on Sidgwick’s understanding of deontology as it commonly appears in lived experience, and on the importance that Sidgwick considers this method to have for ordinary moral reasonings. The ‘distinguishing characteristic’ of Intuitionism, according to Sidgwick, is its belief in ‘certain kinds of actions...unconditionally prescribed, without regard to ulterior consequences’ – and these actions admit of being perceived, with some degree of immediacy, by all people in their general, day to day moral thinking. This is where Sidgwick perceives the practical value of Intuitionism to lie, as it stands. It may be imperfect, and in need of systemisation – but it is a clear and important point of moral reference for all ordinary people.

2.3.b. The Question of Self-Evident Principles: Sidgwick’s Critique of Intuitionism

Sidgwick is nothing in *ME* if not concerned to ascertain what can truly be called reasonable, or *right to do*, and *ME*’s third book is unambiguously written in the spirit of that aim. Having stated that common sense morality, or Dogmatic Intuitionism, is

³⁹¹ When Sidgwick mentions the ‘loose combination or confusion of methods’ that he proposes best represents the moral reasonings of ordinary individuals (I:VIII:102) – it is with surprising indifference considering that this will also be at the heart of his own philosophical moral theory. It is true that in the context of this chapter Sidgwick is mainly making an observation about how morality *is*, rather than how it ought to be – but he then continues on to outline that ‘most moral men believe that their moral sense or instinct in any case will guide them fairly right...still for systematic direction of conduct, we require to know on what judgements we are to rely as ultimately valid’ (I:VIII:103). The latter part of this sentence is actually the starting point of Sidgwick’s own investigation into how valid the precepts of common sense morality really are, from which he will eventually draw the conclusions that make up his moral theory (see also III:I:211-215). There is a parallel here then, between how Sidgwick perceives ordinary morality ‘to go’, and the normative theory that he will later develop.

important in its way, it is now Sidgwick's aim to assess how far the Intuitional system stands up to real scrutiny as an actual system of practical moral guidance. Sidgwick speaks of 'current opinions to which familiarity has given an illusory air of self-evidence' (III:I:212), the 'discrepancies between the judgements of different individuals' (III:I:214), and emphasises the need to 'eliminate error from our moral intuitions' (III:I:214). It is clear that for Sidgwick the enunciations of Intuitionism are going to have to meet an exacting standard³⁹². His conclusions on whether they achieve this are not, for the most part, positive, but understanding the reasons for this gives in turn the grounds for understanding Sidgwick's later re-introduction of his own concept of duty. For Sidgwick will not find all deontological principles to be either invalid, or merely subordinate to another principle, such as utilitarianism. It is a combination of these two conclusions on common sense morality that frame Sidgwick's own use of a deontological version of moral duty.

In III:I, Sidgwick begins laying the groundwork that will justify the meticulous and protracted investigation of the Intuitional method that occupies the next eleven chapters of *ME*. He starts with precisely how to identify the moral value, or the formal rightness, of an action (III:I:202-207) – and concludes that notions of rightness must be limited to what is *objectively* right³⁹³. In this context Sidgwick puts forward one very well known standard of objective rightness - Kant's central axiom of 'Act as if the maxim of thy action were to become...universal law' – and refutes it.

It is important to note from the outset here that Sidgwick does recognise the importance of Kant's rule. He draws attention to the difficult problem of objective versus subjective rightness, and admits that this is where a maxim such as this is needed, in order that it can 'disperse the false appearance of rightness which our strong inclination has given to it' (III:I:209). So effective is this test in fact, Sidgwick

³⁹² This is not the same thing, Sidgwick points out, as asking whether or not these intuitions actually exist (III:I:211). This confusion arises when 'intuition' is taken to mean that such a perception is actually true. Sidgwick is clear that this is not what he means by the term, and that he only intends it to represent the idea that it is known immediately, and not as a result of reasoning. This statement may appear to conflict with the point made above that for Sidgwick, Reason and Intuitionism are inherently linked by the fact that Reason incurs the 'ought', which makes unconditional duties out of moral actions – the duties that characterise the Intuitional method. But the crucial difference lies in the fact that that earlier discussion of ought appeared in the context of Sidgwick's normative thought, whereas here he is conducting an investigation into what *is*. At the start of this investigation, all 'intuitions' are in question. Only those that *do* prove themselves via a process of reasoning do become oughts. Sidgwick states himself on p. 211 that he is at this stage currently holding all moral intuitions 'commonly so called' to possibly be erroneous, or illusory.

³⁹³ Subjective rightness is important to a degree – but, as Sidgwick says, 'too simple to admit of systematic development' (III:I:208).

observes, that Kant elevated it to the point at which it provides the rule for all other rules. But Sidgwick undoes Kant's argument for the the maxim with brutal simplicity. The error lies, Sidgwick says, in mistaking 'Formal Logic' as being capable of providing 'a complete criterion of truth'. It is not that the test is wrong – certainly a rule which does not pass the test is not valid (III:I:210). But it does not necessarily follow that all rules that do stand the test are right. This is because it is simply not enough that any one individual truly *wills* that his action in this circumstance be universally accepted as the course of action for any similar person in any similar circumstances. It is perfectly possible, Sidgwick argues, for an individual to genuinely will that their action become universal law, while the action is actually morally reprehensible³⁹⁴. Sidgwick therefore refutes the maxim on its own grounds. Rather than it solving the problem of gap between objective and subjective rightness, it 'obliterates' that distinction altogether, creating in its place a situation in which whatever someone thinks to be right, *is* so: and as Sidgwick says, '...such an affirmation is in flagrant conflict with common sense; and would render the construction of a scientific code of morality futile: as the very object of such a code is to supply a standard for rectifying men's divergent opinions' (III:I:210)³⁹⁵.

Here, then, Sidgwick has parted ways from Kant in one of the most fundamental ways possible³⁹⁶. This is not on all counts a complete surprise –

³⁹⁴ Sidgwick reinforces his point with the observation that what people believe to be subjectively right, they are also highly likely to believe to be objectively right also. This is basically what is at the heart of the 'conscientious disagreements' that we 'continually find' among people, by which Sidgwick was inspired to write *ME* in the first place (III:I:210).

³⁹⁵ We should note that despite his simple yet effective challenge to the Kantian maxim, Sidgwick does believe that moral intuitions really exist. Sidgwick is himself an intuitionist to a recognisable extent (this will be made explicitly clear by the PD, where it will also be shown, as per my overall argument, that Sidgwick is an Intuitionist (of the deontological variety) on a very specific, and utilitarian, basis). The rules that that intuition produces may not be, in their rawest and most immediate form, ultimately self-evident/reasonable and therefore absolutely true as such, but Sidgwick is not relinquishing his view that there are some moral premises that are intuitively known to be right, independent of consequences. Even when Sidgwick brings this idea together with his views on the role of the utilitarian principle, and develops his own intuitions that *are* self-evident, that basis of Intuitionism remains – it is not simply swallowed or overwritten by the utilitarian principle. It is for the synthesis to explain this in full in chapter 4, but it is worth noting at this stage for Sidgwick's continued valuing of deontological-based principles.

³⁹⁶ Sidgwick includes a footnote on p. 210, stating that he is not prepared to accept Kant's maxim in this form, but that he will include qualifications for accepting it in another form later. This particular instance proves the importance of the PD, and also demonstrates part of the reason that Sidgwick felt compelled to write it. The addition of the PD is thus part of what makes it possible to argue for the strong deontological undertone in *ME* – but it must also be borne in mind that in the PD, Sidgwick is explaining the process of thought through which he went *before* writing *ME*. The PD is not evidence of Sidgwick having 'changed his mind', or come to a clearer conclusion *since* writing *ME* – it is evidence of what was always included in that volume from the beginning. This adds considerable legitimacy to

presumably if Sidgwick had thought that Kant had had it entirely right then this would have been the angle of *ME* from the start. But it does seem to now throw the status of the traditional deontological position in *ME* into question. If Kant's categorical imperative does not yield trustworthy results, then what method can be used to ascertain unconditional, categorical duties? This is, in fact, if those duties can be ultimately established at all.

As I have elsewhere indicated, Sidgwick will actually revisit his (negatively stated) version of Kant's maxim, and it will, in a way very similar to Kant's, yield the sort of result that Kant originally intended, in Sidgwick's own theory. Meanwhile, Sidgwick now asks that very question of how the property of objective rightness is really to be confirmed. The language Sidgwick is using has shifted, away from the more general notion of 'validity', to precisely what is meant by that term: Self-evidence (III:I:212, 213)³⁹⁷. Sidgwick returns here to the apparent agreement on moral rules that is present in common sense morality. But, as demonstrated by the reasons for the fallibility of Kant's maxim, the process of determining validity must be, as far as possible, agent-neutral. The only way then of eliminating error from our moral intuitions is to be found in examining these general rules which Sidgwick calls the morality of common sense, via a more *scientific* approach³⁹⁸. It is only through this sort of approach that real moral duties can be identified, and through which any consistency might ever be reached³⁹⁹.

the argument that Sidgwick had always had the value of deontology in mind, even as he began writing in the late 1860s, while still in his twenties.

³⁹⁷ Sidgwick does not give an exact definition of 'self-evidence' in this chapter – this he reserves for III:XI, in which he specifically tests the maxims of common sense morality against a criteria for self-evidence. It can be inferred however that by 'self-evident', Sidgwick means something that can be found to be ultimately reasonable and absolutely true, or to have the property of 'ultimate certainty' (III:I:214), independently, and without the need for further reference to any other source, definition, or results, for its qualification and justification (see also III:I:200, and III:XI:337).

³⁹⁸ It is worth quoting Sidgwick in full here: "Hence it seems that if the formulae of Intuitive Morality are really to serve as scientific axioms, and to be available in clear and cogent demonstrations, they must first be raised – by an effort of reflection which ordinary persons will not make – to a higher degree of precision than attached to them in the common thought and discourse of mankind in general. We have, in fact, to take up the attempt that Socrates initiated, and endeavour to define satisfactorily the general notions of duty and virtue which we all in common use for awarding approbation or disapprobation to conduct' (III:I:215).

³⁹⁹ Sidgwick once again takes care to state here, at the end of these observations about the fallibility of Intuitionism, that he is 'not trying to prove or disprove Intuitionism, but merely by reflection...to obtain as explicit, exact, and coherent a statement as possible of its fundamental rules' (III:I:216). This is the last time he will make such a statement however: The aim has now unmistakably shifted from being an investigation into what people *do* believe, towards an investigation into what people *ought* to believe. I make the case for Sidgwick's move from his originally intended neutrality to his normative argument in chapter 4.

From III:II to III:X, Sidgwick is engaged in the somewhat laborious process of identifying the commonly received dictates of common sense morality as exactly as possible, and examining each in turn in order to ascertain the degree to which they admit of being ultimately valid, or self-evident⁴⁰⁰. In III:XI, Sidgwick introduces the four conditions that he argues would establish a self evident principle (III:XI:338-342), and invites the reader to consider with him how far those principles imparted through common sense morality can be 'classed as Intuitive Truths' (III:XI:343)⁴⁰¹. From pp. 344 to 355, Sidgwick uses the criteria to renounce one Intuitionist maxim after another, denying that they are absolute rules of duty, or self-evident axioms. As soon as we attempt to define them as clearly as the scientific method requires, we find that there are too many alternative views, too many instances in which a definite rule cannot be obtained, too many cases in which the result is so complicated that any vestige of self-evidence vanishes (III: XI:342-343). In many cases, Sidgwick argues, the principles actually only really point to utilitarianism (III:XI: 342, 348, 349, 350, 352, 354)⁴⁰². Finally, most clearly fail on the fourth criterion, the universality of acceptance⁴⁰³. If nothing else, this has been Sidgwick's

⁴⁰⁰ Sidgwick claims at III:XI:338 that he has only been endeavouring to ascertain which empirically are the general principles or maxims in common sense morality, not necessarily whether or not they *should* be maxims. He has not presumed that just because the moral opinions of ordinary people are 'loose and shifting' that it necessarily follows that no precise principles can be uncovered, and he takes great care once again to emphasise that he has not introduced his own view at any point during that discussion. This could be read as a very fair approach to the investigation, given that Sidgwick has not begun his search from any criteria already imposed by himself. However, it appears evident from the text that his claim on p. 338 is not entirely true. He finds wisdom and self-control, for example, to be only means to the ulterior end of the 'performance of Duty'; Laws and Promises are only binding on the basis that certain other conditions are fulfilled; Liberality and Generosity actually submit to the still greater end of Benevolence. He has been both identifying the maxims of common sense morality, *and*, in some cases, passing judgement on them. In any case, in III:XI he actually introduces the criteria for self-evidence.

⁴⁰¹ These criteria are a) the principle must be clear and precise, b) its self-evidence must be ascertained by careful reflection, c) the propositions accepted as self-evident must be mutually consistent and d) if the proposition is to be considered true, it cannot be capable of being denied by another; that is, it must be universally accepted.

⁴⁰² Sidgwick does theorise throughout these passages that utilitarianism has sometimes appeared to be the more ultimate principle. But he is (at this stage) only observing that the utilitarian formula be admitted to be a 'clear and generally accepted principle' (III:XI:348). He does not state the utilitarian principle to be the answer, only that whereas common sense morality obviously does not accept the utilitarian principle (because the utilitarian principle relies on consequences, which Sidgwick has argued are not relevant to Intuitive principles), it does not definitely affirm any other (III:XI:349). See also IV:III, where Sidgwick talks about the 'latent Utilitarianism of common sense' (IV:III:438). I argue however that the angle of Book IV needs to be read in the context of his wider argument for the negative and positive proof of utilitarianism, that emerges from Book III, and in which it is seen that utilitarianism and common sense morality are not actually juxtaposed. This is addressed in the next chapter (on Sidgwick's utilitarianism), and in chapter 4.

⁴⁰³ Regarding the demand for universal acceptance, Sidgwick experiences something of a quandary. He repeats his earlier point that something that is accepted as universal cannot be held on this

point all along in *ME* about the limitations of the Intuitional position to provide a complete system of morality; there are simply too many interpretations, too many apparent reasonings.

Sidgwick's conclusions on common sense morality that he gives at the end of III:XI are clear. He believes that as no proposition can be elicited from common sense morality that admits of being a scientific axiom, this method of ethics has reached the limits of what it can contribute to a systematic co-ordination of moral principles. The Intuitional method, that aspect of human morality which apparently identifies unconditional, categorical and non-consequentialist duties, is in need of 'regulation and limitation which it cannot itself supply' (III:IX:343). It still has a role; Sidgwick is very specific that he is *not* denying that we have these distinct moral impulses, and that they have a certain authority to which there is a rough agreement⁴⁰⁴. But they are inadequate – or even problematic - by themselves. 'In short', Sidgwick says, 'the Morality of Common Sense may still be perfectly adequate to give practical guidance to common people in common circumstances: but the attempt to elevate it into a system of Intuitional Ethics brings its inevitable imperfections into prominence without helping us to remove them' (III:XI:361).

2.4 The Deontological Nature of Justice, Prudence, and Universal Benevolence

The deontological position (i.e. the system of morality that holds to unconditional and categorically prescribed duties that are ultimately right without reference to consequences, or anything further) has not, then, fared well under Sidgwick's scrutiny. The maxims of common sense morality certainly cannot be understood to be duties in an absolute, inviolable, Kantian sense. Even more gravely, they cannot be called duties according to Sidgwick's own definition of moral obligation. In both cases, the reason is that they are not self-evident, unconditional, or right without reference to anything further. As their 'rightness' is not established by what is *ultimately rational*, they do not incur the 'ought' which would make them unconditional, and categorical.

ground alone to be true, yet he recognises that we allow this weight of agreement in the empirical sciences, so agreement must in some form be an 'indispensable negative condition of the certainty of our beliefs' (III:XI:342).

⁴⁰⁴ Later, Sidgwick actually uses the word 'incontrovertible' to refer to the principles of 'current moral opinion' (III:XIII:379). See below.

In the wake of this inquiry and its results – and the fact that the utilitarian principle has been alluded to as the possible systemiser - it is understandable why Sidgwick is so universally considered to be a utilitarian. But it should immediately be noted here that Sidgwick is only contesting the maxims of common sense morality that are commonly held to be duties. He is not necessarily contesting the *concept* of non-consequentialist duty itself – only that these maxims thus held and expressed do not admit of it. This is to say that Sidgwick’s assault on the non-consequentialist duties of common sense morality does not mean that this particular concept of duty does not have a place in Sidgwick’s work – and it is the remit of this section to now argue that it does. As Sidgwick is now asking whether there really is any possibility of ‘attaining...real ethical axioms – intuitive propositions and real clearness and certainty’ (III:XIII:373), III:XIII marks the beginning of Sidgwick’s development of his own moral theory, and as such the point at which I argue he begins to draw out the aspect of that theory that relies on this deontological version of moral obligation. Furthermore, it will be seen that his moral theory actually requires that this foundation of duty is established *first*. That is, the first time that Sidgwick actually concludes on his utilitarian principle, it occurs as a necessary aspect of his use of deontological concepts. This is due to the fact that Sidgwick’s own ultimate principle of Universal Benevolence, which is at the core of his utilitarianism (see chapter 3) – is informed by self-evident, non-consequentialist, deontological principles. Chapter 4 will expound the actual *function* of these principles within Sidgwick’s moral theory. Here, it is their *nature* that is important, appearing as they do through a process of reasoning, in exact accordance with Sidgwick’s own definitions of moral duty as unconditional, universal, and right without reference to ulterior results, that were given at the start of this chapter.

2.4.a. Philosophical Intuitionism

In III:XIII, Sidgwick returns to the Philosophical Intuitionism that he predicted in I:VIII would eventually be required if the morality of common sense is to be accepted⁴⁰⁵. This will provide the ‘philosophic basis’ through which we might obtain principles

⁴⁰⁵ ‘Is there, then’, Sidgwick asks, ‘no possibility of attaining, by a more profound and discriminating examination of our common moral thought, to real ethical axioms – intuitive propositions of real clearness and clarity?’ (III:XIII:373)

which can be called absolutely, undeniably, and evidently true (see I:VIII:102)⁴⁰⁶. It is important to the argument for Sidgwick's deontology that he maintains this faith in the role of non-consequentialist, absolute-principle intuitions. Despite the disappointing results of his investigation in Book III, Sidgwick is reluctant to simply surrender the method, stating that 'it would be disheartening to have to regard as altogether illusory the strong instinct of Common Sense that points to the existence of such principles' (III:XIII:379). Sidgwick states that the history of moral philosophy shows that philosophers have always been attempting to systemise and correct the common moral thought of mankind, but that 'philosophers have too easily been led to satisfy themselves with ethical formulae which implicitly accept the morality of Common Sense *en bloc*, ignoring its defects" (III:XIII:374)⁴⁰⁷. But Sidgwick's confidence in his own ability to get common sense morality to yield ethical axioms is apparently greater – and the statement he makes to this effect is even more revealing of the importance Sidgwick attaches to common sense morality (or the deontological properties therein): 'We shall find, I think', he says, 'that the truth lies between these two conclusions. There are certain absolute practical principles, the truth of which, when they are explicitly stated, is manifest; but they are of too abstract a nature, and too universal in their scope, to enable us to ascertain by immediate application of them what we ought to do in any particular case; particular duties have still to be determined by some other method' (III:XIII:379)⁴⁰⁸.

⁴⁰⁶ At a practical level, Sidgwick clearly rejects Dogmatic Intuition as a reliable form of moral epistemology on the basis that no fixed or absolute rules can be derived from an intuitional method by itself. It would seem as though Sidgwick agrees with Perceptual intuitionism, given he allows that immediate and general moral principles can be perceived like this – but these principles remain non-inferential. (See also Audi, *The Good in the Right*, p.6).

⁴⁰⁷ Caution needs to be applied, he says, by a mind that is 'earnestly seeking for a philosophical synthesis of practical rules' – there is a need to be wary of principles that appear to be certain and self-evident, but are actually just tautological (something Sidgwick rather entertainingly refers to as 'sham-axioms'). Sidgwick also swiftly rebuts all other attempts to systemise and understand moral thought (III:XIII:377-379). Common sense morality may have continually deceived philosophers into believing that they have secured universal principles, but it is plain to Sidgwick that this is not the case. The principles remain 'incontrovertible, but tautological and insignificant' (III:XIII:379). There is an interesting juxtaposition here between the words 'incontrovertible' and 'insignificant', which is indicative of Sidgwick's emergent idea that the deontological principles of common sense morality are irrefutable in their existence, yet in need of something else to give them real weight, applicability, or meaning. For the argument that Sidgwick does attribute value to the Intuitional/deontological method of ethics, the word 'incontrovertible' is imperative.

⁴⁰⁸ This is the first of the two passages which I refer to as 'the connection passages'. The great significance of these two passages when taken together is fully addressed in thesis chapters 3 and 4. Here, this first connection passage provides the earliest stage of Sidgwick's own moral theory, and a clear indication for my argument that that moral theory is comprised of an amalgamation of deontological and utilitarian principles.

It is widely held that Sidgwick's own philosophical intuitionism is utilitarian – that the principles he establishes as truly self-evident axioms simply serve the purpose of leading to his ultimately utilitarian moral theory – and the passage just cited is deployed as evidence of that view. It would certainly seem to support it when it is thought that the 'some other method' is the utilitarian principle. I do not dispute that the principle (Universal Benevolence) Sidgwick will derive from his philosophical intuitions is in many ways utilitarian in nature – but I *do* argue that the true significance of that passage lies not in its last words, but in its first – 'the truth lies between these two conclusions; There are certain absolute practical principles...'. This indicates that whereas Sidgwick will eventually derive a utilitarian principle from philosophical intuitions (as I am about to expound), those self-evident philosophical intuitions, which are developed out of Sidgwick's discussion of common sense morality, are actually deontological in nature – and they are *prior* to any other principle. In the following section, I argue for this first of the 'two conclusions' - the deontological nature of the self-evident philosophical intuitions of Justice and Prudence - upon which Sidgwick then builds his moral theory of Universal Benevolence.

2.4.b. Justice and Prudence as Deontological Principles

Immediately following his statement about the two conclusions, Sidgwick begins developing the self-evident principles that will inform his own normative moral theory – and for this he goes straight back to the Kantian maxim as the starting point. It was seen above that Sidgwick has already alluded to this principle at I:III:33, where it appeared in the context of his agreement with Kant that 'ought' directly relates to the recognition of rational action. In order for the principle to be made truly self-evident however, Sidgwick argues, it must be stated negatively:

'it cannot be right for A to treat B in a manner in which it would be wrong for B to treat A, merely on the ground that they are two different individuals, and without there being any difference between the natures or circumstances of the two which can be stated as a reasonable ground for difference of treatment' (III:XIII:380) ⁴⁰⁹

⁴⁰⁹ This maxim, Sidgwick admits, is still not complete for guidance, given that an agent must justify a treatment applied to another of which he would complain if applied to himself (i.e. agents may still have to refer to a further rule). But the terms of statement itself, Sidgwick believes, *are* self-evident, in that it itself does not need to refer to any further rule for its truth.

This is the agent-neutral principle of Justice⁴¹⁰. And it is, I argue, is a non-consequentialist maxim. It is self-evident, and therefore ultimately rational, and therefore *right* in Sidgwick's sense of 'not fit to an ulterior end'. i.e. the maxim is not simply 'right' with regard to ulterior results; it is wholly self-referencing, and right *independently* of anything further. Also, given that it is ultimately rational, it incurs – as per Sidgwick's earlier argument for the motivating force of reason – the normative 'ought'. It therefore becomes a categorical imperative, again as per Sidgwick's earlier arguments for 'ought' and 'categorically prescribed', and it is unconditional⁴¹¹. Furthermore, the maxim has the element of universalizability. This is not an aspect of traditional deontological thought that has featured much in Sidgwick so far, but it is clearly present here. Not only then does Sidgwick's principle of Justice retain all of the vital characteristics of Sidgwick's own definitions of non-consequentialist rightness and moral obligation, but it acquires a *further* deontological characteristic, when stated by Sidgwick in this way. Where Sidgwick has pointed out previously that Kant's maxim is not an effective method of establishing truly objective and universal duties, Sidgwick's negative version of that maxim *does* do this. Here, Sidgwick is almost more Kant than Kant. These qualities combined render the principle of

⁴¹⁰ There are interesting implications here for the use of Sidgwick and Kant respectively in politics, and especially for the topic of social contract theory, and the discussion so prevalent in that area of what exactly is meant by 'liberty'. Where Kant's positive maxim seems to further the case that liberty is to be understood as 'freedom to', i.e. as the basis of a rights-based version of liberty, the sort taken up by Rawls in his neo-Kantian work on rights, Sidgwick's maxim supports the argument that liberty represents 'freedom *from*' certain treatments. Both interpretations of liberty are essential to a full understanding and application of that concept, in much the same way as both deontological and utilitarian principles of morality are vital to a full understanding of human morality, especially at the level of practical ethics. In both situations, it would seem as though a moderate view is needed of each interpretation – that is, a balance is required between their respective components. In this way, the central argument of this thesis – that morality is comprised of a delicate balance between two apparently conflicting yet simultaneously vital principles – has significant potential for extending into political philosophy.

⁴¹¹ The question arises here as to whether Sidgwick successfully avoids the problem that he sees in Kant's maxim, of it obliterating the distinction between objective and subjective rightness by allowing that whatsoever a person *genuinely* wills become moral law does so, even though he may, objectively speaking, actually be in error (see III:I:210). The answer seems to be that Sidgwick does avoid this problem, mainly because the negative version is more obviously self-evident in that it cannot be any further proven without necessarily referencing itself. It could also be argued that Sidgwick avoids the objective/subjective problem here by way of the simple fact that he does not prescribe a rule for judging actions themselves, but a broader theory that gives the framework, or terms, that is to be prior to any action. This suggests that Sidgwick's maxim requires expansion if it is to be used as a guide for action – something that Sidgwick himself recognises when he states that his principle 'manifestly does not give complete guidance'.

Justice a recognisably deontological duty – and this categorical and unconditional duty is one of the bases from which Sidgwick's whole moral theory proceeds⁴¹².

The second self-evident principle at which Sidgwick arrives, directly after having established Justice, is that of Prudence. In the same way as the principle of Justice (or 'equity', as Sidgwick describes it here for emphasis of his next point) was obtained by considering the similarity of all individuals that make up the whole, so too must what is *good* for an individual be comprised of all the goods, or concerns, that make up our 'conscious life' (III:XIII:381)⁴¹³. This includes all future good, as well as present ones; thus Prudence represents the notion of temporal impartiality. This, Sidgwick believes, is also ultimately self-evident, and reasonable, for one's own good is conceived as a mathematical whole, 'of which the integrant parts are realised in different parts of moments of a lifetime', and to disregard or sacrifice other parts of that whole is *unreasonable*.

'...mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment than to that of another. The form in which it practically presents itself to most men is 'that a smaller present good is not to be preferred to a greater future good'...' (III:XIII:381)

As such, the impartiality principle of Prudence is also, like Justice, a non-consequentialist principle: It is ultimately rational, and therefore true without reference to anything outside of itself. In this meta-ethical context, it is thus also a categorical and unconditional duty.

It is in this way then that I argue Sidgwick's philosophical intuitions of Justice and Prudence are recognisably deontological in nature. Given that they are self-evident and thus ultimately reasonable, they are *right* without reference to any consequences that they might bring about. They also, being ultimately reasonable, both incur the motivation to action that the recognition of that rationality involves – and that motivation to action is the same sort of unconditionally required 'ought' that is found in the traditional Kantian deontological school. It is true that Sidgwick will proceed to also attach this kind of ought to a utilitarian principle – but this does not

⁴¹² This is also the building of Sidgwick's own moral theory of Rational Benevolence, but it is the concern of chapter 4 to present this as that theory. Here, the aim is merely to show the predominantly deontological character of even the very origins of Sidgwick's moral theory.

⁴¹³ See also I:IX:111-113

detract from what I argue to be the deontological origins of these, the very foundations of his moral theory.

2.4.c. Universal Benevolence as a Deontological Principle

The last of Sidgwick's philosophical intuitions is that of Universal Benevolence, and in this section I make the case that this too – despite its obvious utilitarian connotations – is deontological to some extent. This is because Universal Benevolence is derived directly from a sequential linking between Justice and Prudence, which I have just argued are recognisably deontological in nature – and because Universal Benevolence is also self-evident. As I said above, it is the remit of chapter 4 to argue for the *function* of this deontology in the establishment of Sidgwick's theory of Rational Benevolence, providing as it does one part of the synthesis between deontological and utilitarian principles that I believe to be integral to that theory. Here however, within the context of this chapter which argues for the significant deontological influence in Sidgwick's thought, a brief overview of how Sidgwick arrives at Universal Benevolence via Justice and Prudence is necessary in order to demonstrate its partially deontological nature.

It will be observed – as Sidgwick does - that the first philosophical intuition of Justice provides only the terms for action, and not the content. For this reason, that maxim cannot itself be considered the kind of duty capable of giving normative guidance for actual action (III:XIII:380). Having then established the second principle, Prudence, according to a similar 'mathematical whole' approach, Sidgwick next argues that as this notion of individual good is logically comprised of all of the goods that an individual experiences, this same integration of all goods must logically occur *between* all similar human beings, as well as within them (III:XIII:382). That is, Justice and Prudence combine to state that as it is irrational to prefer one good over another at the individual level, it is also irrational to prefer the good of one individual over that of a similar other. Ultimately then, if we 'ought' to aim at our own good because this is only rational, then we ought to aim at Universal Good, given that this

too is rational⁴¹⁴. And it is this that yields the ‘point of view of the universe’, or, Sidgwick’s self-evident principle of Universal Benevolence:

‘...the good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other; unless, that is, there are special grounds...and it is evident to me that as a rational being I am bound to aim at good generally...and not merely at a particular part of it’ (III:XIII:382)

Sidgwick considers this Universal Benevolence – which states that each person is *morally bound* to regard the good of any other individual as much as his own - to be the result of self-evident ‘rational intuitions’⁴¹⁵. Hence the full term for this position, ‘Rational Benevolence’.

Now, we also see the utilitarian character of Benevolence clearly (I will return to this aspect of Benevolence shortly) and the argument that it is also in some sense deontological does require a strong line of defence⁴¹⁶. But it is my view that Sidgwick provides that defence, in his definition of rightness *not* as fit to an ulterior end, which – it will be recalled – applies just as much to ends as it does to certain kinds of actions. It is only because of Sidgwick’s understanding of right as being that which is ultimately rational, and therefore *right without reference to ulterior results*, that Benevolence can be called self-evident at all – once Sidgwick has found it to conform to those requirements. Furthermore, I venture to state that Benevolence is in fact non-consequentialist. As an immediate qualification to this claim, it is not

⁴¹⁴ It is not my concern in this thesis to tackle Sidgwick’s Dualism of Practical Reason, but I do believe that Sidgwick gave better grounds than he thought he had for arguing that the *general* good is to be more rationally preferred than individual good. As far as I can tell, Rational Benevolence is the best example of this.

⁴¹⁵ Sidgwick states that this Benevolence is a ‘necessary inference’ - an efficient alternative way of claiming that something is self-evident – an assertion that Sidgwick confirms a few pages later when he describes it as ‘an indubitable intuition of the practical Reason’ (III:XIV:400). Sidgwick admits that previously, in his investigation into the precepts of common sense morality, Benevolence was not found to be self-evident principle. But this may be overcome, he argues, with the inclusion of the notion that people *ought* to aim for the good of as many as are closely connected with them. That is, when Benevolence was discussed previously, it was within the context of Sidgwick’s observations about the nature of the principles derived from common sense morality *as they stand*, and not within the context of Sidgwick’s argument about which principles people *ought* to obey. See also III:XIII:387, where Sidgwick refers to Rational Benevolence as an axiom, rather than just a maxim.

⁴¹⁶ As discussed in chapter 1 (pp.130-134) of this thesis, it is precisely the ‘point of view of the universe’ passages that for Peter Singer prove beyond question Sidgwick’s status as a hedonistic utilitarian. The utilitarian quality of Universal Benevolence is further explored both in chapter 3, which specifically examines Sidgwick’s utilitarianism, and chapter 4, where this point that both deontological and utilitarian properties are essential to Universal Benevolence is the central argument.

meant that Benevolence is ‘non-consequentialist’ in terms of its dictates; it is a utilitarian principle after all, and it will be important to my overall argument for Rational Benevolence that it is. But as an *end*, an ultimate rational end, the substance of the principle itself is non-consequentialist, in that it is ‘right’ according to Sidgwick’s definition as right without regard for ulterior results. We can certainly say that Benevolence is unconditional, in that – again - it is not contingent on anything further, and that as a rational end it becomes a categorical imperative, and thus incurs the normative ought. It is my view that on these grounds, Benevolence can be understood to be, at least in part, a deontological kind of moral duty.

Finally, I will add that as the foundations, or the *basis*, of Benevolence – i.e. the principles of Justice and Prudence - are deontological, Benevolence remains partly deontological in this way also. Sidgwick is clear in *ME* that the utilitarian principle is in need of a proper basis; Justice and Prudence are the cornerstones of that basis - they cannot simply be ignored once Universal Benevolence has been derived, or Universal Benevolence will once again be without its proof⁴¹⁷. The deontological status of Rational Benevolence is thus doubly supported – once by having its foundation in deontological principles, and again through its own characteristics of being ultimately rational, unconditional, and right without regard to ulterior results.

2.4.d. Rational Benevolence as a Deontological Duty

It remains now to assess how far we can call Universal Benevolence a deontological duty in a practical sense. It is my view that in Sidgwick’s argument for Rational Benevolence, we see the convergence of his various points made throughout *ME*, now cohering to create his fully normative theory – and, as I have argued in this chapter, many of these points had recognisably deontological elements. There was his theory of moral obligation, in which Sidgwick’s interpretation of ‘right’, or ‘ought’ is that it denotes something that is categorically required, purely for the sake of its own properties ; there was his recognition of the value of the Intuitionist view, defined as

⁴¹⁷This idea that the deontological properties of Justice and Prudence are the ‘proof’ of utilitarianism is essential to my argument for the synthesis between deontological and utilitarian principles, and is expanded upon fully from the end of chapter 3, and throughout chapter 4.

the non-consequentialist-duty position, and the common sense morality therein; there was the fact that Sidgwick used as the point of departure for his own moral theory the search for the self-evident principles which the Intuitionist view believes itself to hold; there is the fact that Sidgwick's self-evident principles are ultimately reasonable, and thus emerge as 'right' in and of themselves; there is, of course, Sidgwick's recognition of the Kantian maxim. All of these points have combined to produce Rational Benevolence, and as the theory of Rational Benevolence also includes the very normative ought itself which was shown to have deontological qualities, it is also normatively deontological. As Sidgwick says at III:XIII:382, we are 'morally bound' to Benevolence: The recognition of this end as ultimately rational brings with it the rational impulse and moral obligation to act upon it. But perhaps the best evidence that Rational Benevolence can be understood as a deontological sort of duty comes from Sidgwick's direct comparison between Rational Benevolence and Kant's 'fundamental principle of duty' (III:XIII:385-386), which he believes to 'coincide to a considerable extent, if not completely', with his own moral theory. 'We find', Sidgwick says, 'that when he comes to consider ends....the only really ultimate end which he lays down is the object of Rational Benevolence...he regards it as evident *a priori* that each man as a rational agent is bound to aim at the happiness of other men...Kant's conclusion appears to agree to a great extent with the view of the duty of Rational Benevolence that I have given'⁴¹⁸.

Lastly on this point, utilitarianism has been conspicuous by its absence so far in *ME*. Whereas this does not in itself necessarily mean that Sidgwick's moral theory has to be deontological, it at least adds to the impression that Sidgwick did not

⁴¹⁸ It is extremely illuminating that it is just after his establishment of Rational Benevolence that Sidgwick chooses to reference Kant more comprehensively than anywhere else in *ME*. On account of Sidgwick's also having found Prudence, or Rational Egoism, to be a self-evident principle, Sidgwick is forced to admit here that he actually doesn't agree entirely with the last part of that passage: Sidgwick agrees, under the influence of Butler, that 'one's own happiness is manifest obligation'. This, of course, leads to the Dualism of Practical Reason with which Sidgwick will struggle in the closing stages of *ME*. But here, Sidgwick points to the positive side of the statement – the *duty* of each person – and concludes that Kant's view seems to agree to a great extent with the view of the duty of Rational Benevolence. These passages can be read as Sidgwick arguing that his theory supersedes Kant's, but this is of no consequence. What is important is Sidgwick's readiness to align his position with that of Kant – and even, if it is thought that Sidgwick is arguing for his superiority, to suggest that his own theory actually explains Kant's theory of duty more effectively. It should be noted here that it is clear Sidgwick does not believe Kant's original maxim – 'act on a maxim that one can will to be law universal' to be the self-evident version of this maxim, and that thus phrased this is actually a *corollary* of the principle of Justice as stated by Sidgwick. But this does not affect the deontological quality of that principle of Justice itself.

simply *begin* that theory from a utilitarian position. It is Sidgwick's methodical approach to 'the methods of ethics' which in the end yields his own normative theory, and he draws the conclusions for that moral theory in part from the deontological system which he has been examining in this systematic way. Again, I emphasise that I am not denying the utilitarian properties of Rational Benevolence – it *is* also a utilitarian duty. But I maintain that the '*duty*' here, that which makes Rational Benevolence a moral obligation, is closer to the traditional deontological interpretation of duty. As to how it can *also* be utilitarian, this is left to the remit of this thesis' chapters 3 and 4, which explore Sidgwick's utilitarianism and argue for the synthesis between the deontological and utilitarian properties of Rational Benevolence respectively.

2.5. Sidgwick's Intuitionism, and Sidgwick's intuitionism

In the previous section, I argued that there is a distinct non-consequentialist aspect to Sidgwick's moral theory. It will also be recalled from earlier in the chapter that non-consequentialism was a part of Sidgwick's definition of Intuitionism (understood as the deontological system of ethics). These, then, are the grounds for my final claim in this chapter about Sidgwick's deontology, which is that part of Sidgwick's understanding of Intuitionism actually appears in his own epistemological intuitionism. This means arguing that Sidgwick's epistemic intuitionism is in fact to some extent deontological – and this accords directly with my argument just made for the deontological character of the rational intuitions that make up Rational Benevolence. But a further and even more specific link can be made between the two types – a link that further proves the presence of a deontological element in *ME*. This link comes in the form of *Kant's* appearance in Sidgwick's intuitionism.

Sidgwick is not often explicit about the Kantian influence in *ME*, even during his discussion of Reason. He is, however, explicit about it in '*Professor Sidgwick's Account of the Development of his Ethical View*' – the autobiographical fragment once included in the preface to the sixth edition of *ME*, and referred to here as the Personal Document (PD). Here, Sidgwick describes the progression of thought through which he passed as he was searching for a viable moral position. At the early point at which he was struggling with proving that it is the general happiness that is paramount, Sidgwick finally concluded that we must simply see that the

general happiness is more important than our own. That is, Sidgwick found himself ‘forced to recognise the need of a fundamental ethical intuition’ (*ME*: xvi). From this point he describes how he then moved directly back to Kant, where he was ‘impressed with the truth and importance’ of Kant’s axiom, ‘act from a principle or maxim that you can will to be a universal law’ (*ME*: xvii). This is a self-evident principle, that does not refer to consequences. A few paragraphs later, Sidgwick again references the notion of ‘a fundamental moral intuition’, this time stating that such an intuition is ‘binding’ (*ME*: xix) ⁴¹⁹. On the final page, Sidgwick again refers to his acceptance of the Kantian principle (*ME*: xx), and describes how *this* principle validated and supported the other fundamental intuition, which is the utilitarian principle of aiming at the general happiness. That Kantian principle to which Sidgwick is referring is Kant’s formulation of the categorical imperative (which in Sidgwick’s hands became the principle of Justice). This form of duty, Kant’s fundamental maxim, had ‘commended itself’ to Sidgwick’s Reason (*ME*: xvii), and it endures even when the utilitarian principle is added.

The PD thus clearly demonstrates that for his own intuitionistic moral epistemology, Sidgwick uses a categorical and unconditional principle that is right in itself without reference to consequences – that is, that he uses a deontological principle. And it is a Kantian principle, no less⁴²⁰. As this non-consequentialism is

⁴¹⁹The similar phrase ‘intuitively known to be unconditionally binding’ also appears at I:I:3, and in that instance Sidgwick does use it in specific reference to duties, in the same way that he also defines Intuitionism within discussions of ‘Duty’ elsewhere (III:I:200). It could be argued however that given Sidgwick’s own adherence to intuitionism as an epistemological device, this phrase ‘intuitively known to be unconditionally binding’ will eventually, in Sidgwick’s hands, take on a utilitarian form: That is, there is nothing to say that that phrase need apply only to deontological systems of ethics (it is on this basis that Phillips describes Sidgwick’s position as ‘intuitionistic utilitarianism’). I agree that ‘intuitionism’ with a small ‘i’ can be removed from the deontological view in Sidgwick’s work and applied straight to utilitarianism instead (Sidgwick himself makes this very point) – and that the passage on xix seems to suggest that this is the approach Sidgwick has taken. However, I argue, contra Phillips, that Sidgwick’s own intuitionism is entirely utilitarian. Firstly, Sidgwick’s flexibility with intuitionistic terminology only further supports this thesis’ argument that ‘intuitively known’ represents for Sidgwick not *just* the process of reasoning that informs the deontological/Intuitional position, but also the process of reasoning which equally informs the utilitarian position – and thus his intuitionism can be thought of as representing the shared appeal to Reason. Secondly, as it has been argued here, Sidgwick retains an element of the ‘duty’ aspect of Intuitionism in his own moral theory, which maintains a deontological influence in that theory. It is on this basis that I disagree with Phillips’ description of Sidgwick’s intuitionism (unless it is posited in line with this thesis that the ‘intuitionistic’ aspect represents the deontological position, which in Phillips’ justification for the phrase it clearly does not).

⁴²⁰ It seems relevant to ask here whether it has been Kant’s doctrine that Sidgwick has had in mind for the non-consequentialist, Intuitional view all along. It might be an advantage if it was. Kant’s version of duty is paradigmatic of non-consequentialist duty-based systems of ethics, and especially of the duty-based systems of ethics that conflict with consequentialist utilitarian systems. As my ultimate aim is to show that Sidgwick reconciles this divide by maintaining a recognisably deontological element in a

also a fundamental part of how Sidgwick defines the moral-theoretic Intuitional view, Sidgwick's Intuitionism and Sidgwick's intuitionism thus share this factor in common. In order to complete this argument, it is required to be shown that Sidgwick's intuitionism as it appears in his normative moral theory is deontological as well as utilitarian - and I believe that I have made that argument above by arguing that Sidgwick's own philosophical intuitions are not entirely utilitarian⁴²¹. Overall, it is my view that as Sidgwick specifically calls the non-consequentialist method 'Intuitionism', and then himself attributes value to both non-consequentialist and utilitarian intuitions, this has caused some of the long-running confusion over how Sidgwick uses 'intuitionism' variously. In Sidgwick's moral theory, Broad's 'classification' and 'epistemic' types of 'intuitionism', coincide, linked by Sidgwick's use of a non-consequentialist form of duty.

2.6. Conclusions on Sidgwick's Deontology

The purpose of this chapter was to establish that, contrary to the prevailing opinion that Sidgwick is a classical utilitarian, deontological principles have a considerable presence and active *value* in Sidgwick's moral theory. It was observed early on that Sidgwick's thought is very much dominated by *rightness*, as opposed to the good,

theory that also includes utilitarianism, it is important that Sidgwick can meet deontological theory at one of its most influential and renowned sources. It is also the case that the Kantian view of Reason and morality was a prevalent version of deontology at Sidgwick's time, and revealing a Kantian influence in Sidgwick's work would show that Sidgwick recognised its value, despite his 'utilitarian' status. But these points only really indicate that it would be *preferable* that it is Kant's system Sidgwick has in mind, and this does not amount to an argument that it actually *is* Kant's system. (There is also Donagan's argument – outlined above- that the intuitionist position to which Sidgwick is referring actually seems to be Whelwell's). The only place that it becomes evident that Sidgwick has a Kantian 'intuition' in mind is in the PD, which tells us that this was the basis for the principle of Justice. It is therefore neither certain nor in fact necessary that the Intuitional position is Kantian from the beginning. But Sidgwick's non-consequentialist view is certainly Kantian in the end.

⁴²¹ Sidgwick himself supports that argument when he points out that his self-evident principles belong as much to Utilitarianism as they do to Intuitionism III:XIII:386-397.

See also Sidgwick's comment in the PD (xx) that, at the end of the process he has described in which he perceives the lack of opposition between Intuitionism and Utilitarianism, he was '...a Utilitarian again, but on an Intuitional basis'. At the corresponding point in *ME*, by the time Sidgwick reaches his own moral theory and states that his rational intuitions/self-evident principles furnish Utilitarianism as much as they do Intuitional ones, this can already be logically inferred from Sidgwick's arguments. Ironically then, if we track the evolution of Sidgwick's 'Intuitionism' alongside that of his 'intuitionism', it seems as though there is both a clearer distinction between the two types than Phillips allows, *and* an even less defined one. But Phillips' argument for the lack of definition between them draws on the apparent confusion between Intuitionism as a method and intuitionism as an epistemic device, and not, as my argument does here, on the idea that there might actually *be* less distance between them than is commonly thought.

and that this pervades *ME*. In terms of how to define this rightness, for Sidgwick this is not 'fitness to an end', but what is ultimately rational, whether this be a certain kind of action, or an ultimate, rational end. Sidgwick's concept of 'ought', then, pertains to what is right in itself without reference to consequences, and I believe that this is a predominantly deontological understanding of rightness, and of moral obligation. It was also seen that Sidgwick defines Intuitionism as the non-consequentialist view, and that Sidgwick believes this method of ethics to actually have some day-to-day value, in the form of common sense morality. The self-evidence of its principles, however, is only illusory. But self-evidence is of the utmost importance to Sidgwick. Self-evident principles will be part of the truth which lies between two conclusions - the 'certain, absolute, practical principles, the truth of which is manifest'. Sidgwick argues these to be Justice, Prudence, and Benevolence, and all of them are non-consequentialist in that they are crucially informed by Sidgwick's interpretation of rightness as being that which is ultimately rational. i.e. these principles are not simply right with reference to their ulterior results. Thus, Rational Benevolence – which is eventually derived from those philosophical intuitions – is also self-evident and right in that same way. It is in this way that the non-consequentialism of the Intuitionist view appears in Sidgwick's own intuitionism – a point that is strongly supported by Sidgwick's use of a fundamental intuition from Kant – and that traditional deontological concepts can be seen to form a part of Sidgwick's normative moral theory.

With these conclusions thus condensed, we can now ask the question that has been latent throughout this chapter: Is it possible, in any way, to call Sidgwick a deontologist? It is my view that some significant movement can be made towards making this argument. Firstly, if Sidgwick's own intuitionism is fundamentally and crucially informed by an intuition from *Kant*, and if Sidgwick believes Kant's position to be what is a recognisably *deontological* position in that it is a form of moral obligation that depends on duties categorically and unconditionally prescribed, then there is a line of Aristotelian logic that allows us to call Sidgwick a deontologist. In the PD, Sidgwick himself actually declares outright that he had 'become....an Intuitionist to a certain extent...' ⁴²². If it is to be understood, as *ME* strongly suggests,

⁴²² In addition to the discussion above as to whether Kant *is* an influence on Sidgwick's 'Intuitionism', it can also be noted that Sidgwick demonstrates a peculiar reluctance to directly reference Kant, even when discussing the role of motives in morality (Sidgwick simply refers to 'moralists of influence')

that by Intuitionism Sidgwick means the position that attributes value to non-consequentialist duty then in the personal document Sidgwick is declaring *himself* to be a deontologist, in *some* sense. After all, how else do we say what it is *ultimately right* to do? If all morality was contingent on the consequences of actions, Sidgwick would be very unlikely to find the answer to his pressing question about ultimately right conduct, as he perceives that question to present itself.

On the other hand, we must obviously take account of the fact that Sidgwick's view of rightness does allow for ought to apply to means to ends to some extent, which detracts from the non-consequentialist aspect of the traditional deontological understanding of duty. Then of course there is the fact that Rational Benevolence, which has at its basis the principle of aiming for Universal Good, is a utilitarian principle. For these reasons, Sidgwick's theory of moral obligation cannot truly be called *completely* deontological. But I do maintain that it is still deontological *enough*. That is, Sidgwick's particular concept of moral obligation as depending on that which is ultimately rational and not simply fit to an ulterior end makes it unconditional and non-consequential enough that it retains a character much more akin to the deontological understanding than to the utilitarian one. Another way of defending Sidgwick's deontology might be to ask the question from the other side, i.e. 'is Sidgwick's theory of duty ultimately completely utilitarian?' For it is clear that this is entirely not the case. Sidgwick is not arguing for rightness as simply fit to an ulterior end, and his argument that rightness is what is ultimately rational demands that his concept of 'ought' also be detached from the notion of consequences. And where rightness, ought, and duty *are* applicable to ends, this is precisely why he can transcend the boundaries between deontological and utilitarian positions as well as

III:I:204). This admittedly opens up an ambiguity regarding whether or not it is Kant's understanding of moral duty that Sidgwick has in mind in *ME*. The influence of what Schultz refers to as 'Germanism' (*Eye of the Universe*, p.175, by which he means Hegel and Kant) in Sidgwick's work has been noted by both Schultz and J.B. Schneewind, and Schneewind in particular insists that whereas 'Sidgwick himself points out the Kantian affinities of his position, he is by no means a Kantian' (*Sidgwickian Ethics*, p.286). Schneewind's position on Sidgwick's use of Kant is that 'he is deliberately developing a traditional mode of approach to basic axioms. In doing so, he brings out distinctly new possibilities within it'. I agree with Schneewind that Sidgwick cannot strictly be called a 'Kantian'. After all, Sidgwick refutes Kant's own argument for his moral axiom, even though Sidgwick himself adapts and then uses that rule. But this does not mean that Sidgwick cannot be called a *deontologist* (in some sense), and nor does it mean that that deontology is not informed by Kant's doctrine. I believe it is important to note here that Sidgwick did not have the term 'deontology' – which today is so freely associated with Kant – at his disposal. I argue – as I have above – that if Sidgwick *had* had that terminology, the language he uses to describe Intuitionism in these passages soundly suggests that he would have directly associated Kant with a 'deontological' position, understood as involving non-consequentialist, unconditional duties.

he does – as I will argue later⁴²³. Even if Sidgwick's theory is not deontological in an official sense, it is far enough away from a standard utilitarian basis to be able to construct a synthesis between deontological and utilitarian principles of morality.

Overall, we can return to Sidgwick's own summary of this situation, which is that the 'truth lies between two conclusions'. The self-evident (deontological) principles are crucial – but they require something further. Sidgwick's critique of Intuitionism reveals that as important as that method is, it is not adequate for the construction of a full system of ethics; this can only be done via 'some other principle'. This other principle is utilitarianism. So here we must follow Sidgwick's lead in stating that he is only a deontologist on a utilitarian basis. It is now the role of the next chapter to draw out the nature and place of this utilitarianism in Sidgwick's work, and of the final chapter to explain how the two types of moral principle are, in Sidgwick's theory of Rational Benevolence, inherently connected.

⁴²³ It is true that Sidgwick appropriates Kant for his own cause to an extent – as will be seen he *will* be using this version of duty to argue for the inclusion of a utilitarian principle in his moral theory. But this is almost my point; Sidgwick's notion of duty began life as the Kantian version, and it retains much of that character in his own moral theory, which is precisely what allows him to establish that utilitarian principle with some kind of true validity.

CHAPTER 3

BUILDING A SYNTHESIS: SIDGWICK'S ACCOUNT OF UTILITARIANISM

Henry Sidgwick is not always included among those moral philosophers who are considered to have offered complete normative moral theories. Where Sidgwick *is* thought to adhere to a particular moral position, it is almost universally accepted that it is a classically utilitarian one, in the same line as Bentham and Mill. However, in chapter 2 of this thesis I concluded that Sidgwick afforded far greater value to deontological principles of morality than is commonly thought, and I argued on this basis that traditionally deontological principles form an essential part of Sidgwick's normative ethics. This finding is crucial to my wider argument that Sidgwick's theory of Rational Benevolence offers a synthesis between deontological and utilitarian principles. But in order to complete the grounds on which this argument can be made, it is still necessary to also gain a full understanding of Sidgwick's utilitarianism. It is the aim of this chapter therefore to explore Sidgwick's utilitarianism as it appears throughout *ME*. Together, chapters 2 and 3 set up the understandings of Sidgwick's deontology and utilitarianism from which the synthesis between them in Rational Benevolence can be developed.

The utilitarian aspect of Sidgwick's work is not just crucial to this thesis' argument for the synthesis. There is no doubt that Sidgwick is a utilitarian in some capacity, and failure to take account of Sidgwick's utilitarianism would inarguably lead to an unfair and inaccurate representation of *ME*. However, where the majority of the Sidgwick literature has been shown to hold that Sidgwick is a *classical* utilitarian, my argument is that Sidgwick's utilitarianism is only of a very specific type, in that it contains qualities and emphases that are not typical of utilitarianism when standard utilitarianism is understood as holding that actions are only morally right when they conform to certain ends. Textual analysis of *ME* quickly reveals Sidgwick's utilitarianism to be more complex than is generally thought: Where Sidgwick's views on deontology are usually easy to identify, and unfold linearly, his discussions of utilitarianism are often somewhat more obscure, with some of the key concepts appearing in close conjunction – and sometimes even agreement – with concepts usually considered to belong to the deontological school. Even Sidgwick's argument for Happiness as a rational, ultimate end is established on grounds that accord more

closely with deontology. I am not, then, disputing that Sidgwick is a utilitarian of sorts – but I am disputing the opinion that he is simply a traditional utilitarian whose version of the doctrine conforms to all criteria of a standard classical utilitarian position, and I argue that it is the less typical features of his utilitarianism that make it possible, or even necessary, to join his utilitarianism to his deontological principles. Most importantly, I argue that the connection is made possible via their shared appeal to *Reason*. I posit that Sidgwick himself perceived this connection, and it is this relationship that will be drawn out in chapter 4 when it is argued that Rational Benevolence is comprised of that synthesis between the two types of principle. The more specific purpose of this chapter therefore is to draw attention to the areas in which Sidgwick's utilitarian position is particularly *distinctive*, and to where this distinctiveness eventually points to a reliance on deontological principles.

Sidgwick's utilitarianism can be effectively understood via four defining features⁴²⁴. The last three of these provide the stages of the chapter's investigation, but the first point can be made here – and it is that Sidgwick was explicit that he did not intend on writing a defence of utilitarianism. Sidgwick's mission statement in the preface to the first edition of *ME* is the firm assertion that he is only attempting an examination, 'at once expository and critical.....of the different methods of obtaining reasoned convictions as to what ought to be done which are to be found – either

⁴²⁴ All of these allusions to Sidgwick's treatment of utilitarianism, and to the more unusual aspects of that treatment, can be found in outline in the prefaces to the first two editions of *ME*. Also in the preface to the first edition, Sidgwick writes (continuing the next passage cited in the main text below) – 'I have put aside temporarily the urgent need which we all feel of finding and adopting the true method of determining what we ought to do...' (*ME*: vi). The appearance of these passages in the first two prefaces is in itself significant. Firstly, this is Sidgwick unambiguously stating that *ME* was never meant to be an argument for utilitarianism. It could be supposed from the critics' responses to the first edition of *ME*, to which Sidgwick is responding in the second preface, that despite Sidgwick's intention to remain neutral, at least the first edition emerged as unmistakably utilitarian. This might then in turn reinforce the idea that Sidgwick believed utilitarianism to be the most accurate and valuable method of ethics, and that he had, through *ME*'s investigations, found this conclusion to be unavoidable. But Sidgwick stands firm against the critics (namely, Bradley and Collingwood) in the second preface, arguing that *ME* is not simply a defence of utilitarianism and an attack on common sense morality. Sidgwick admits that he has criticised common sense morality 'unsparingly', but that he considers himself to have 'exposed with equal un-reserve the defects and difficulties of the hedonistic method' (x). Even after prolonged accusations of the very sort that Sidgwick had wished to avoid, Sidgwick remains committed to his aim of neutrality. He thus provides his own unequivocal evidence that he was not simply justifying utilitarianism.

Secondly, Sidgwick's comments in the prefaces – the very earliest stage of *ME* - demonstrates Sidgwick's willingness to expose his position on certain important aspects of his approach to utilitarianism even from the outset. It is undoubtedly detrimental to *ME* – and to Sidgwick - that in the current day, *ME* is hardly ever published with those prefaces included. Omitting these crucial pieces of primary evidence for Sidgwick's thought is likely to lead if not to a misreading of Sidgwick, then at least to a less detailed reading, and this is hardly fair.

explicit or implicit....' (*ME*: v). That is, he is *not* aiming to defend any one position; his intention is to remain completely neutral as regards which ethical system he himself believes to provide the best guidance for moral conduct, or to generate the most morally valid results. This is vitally important; it is the first evidence that *ME* does not necessarily present Sidgwick as a utilitarian, and it is Sidgwick himself who makes the statement, both early, and emphatically⁴²⁵.

The second feature is Sidgwick's understanding of rightness, which includes his rejection of rightness being defined as fit to an ulterior end, and his resulting non-utilitarian position on means to ends. This in turn indicates the need to assess Sidgwick's view of Happiness as an ultimate end, and the complex and often conflicting relationship between Sidgwick's doctrine and that of Bentham's. It also calls into question the teleological nature of utilitarianism in Sidgwick's work, and the Benthamite principle of utility itself.

The third point is that Sidgwick draws attention to the 'difficulties and defects' of the utilitarian position (*ME*: x). This particular view of Sidgwick's unfolds as *ME* progresses, with his discussions of utilitarianism rarer and often appearing more ambiguous than those of other methods (egoism and Intuitionism). Despite this, Sidgwick's utilitarianism is highly sophisticated in its depth, and, vitally, it is the proficiency with which Sidgwick handles utilitarian concepts that eventually produces his own final version of the doctrine –one that cannot be fully understood without reference to certain deontological principles.

This leads to the fourth and final point, which is that Sidgwick is eventually brought to a position, in Book IV of *ME*, from which he makes the case for the utilitarian position. But the material here is explored in terms of the question as to how far Sidgwick considers utilitarianism alone to really be capable of serving as a complete and holistic theory of human morality. Crucially, the value that Sidgwick ascribes to utilitarianism appears not entirely in Book IV (which contains Sidgwick's 'positive' proof for utilitarianism), but mainly in Book III, where it emerges in *inherent conjunction* with deontological principles (in the form of his 'negative' proof). In Book IV itself, Sidgwick actually refers the reader back to the conclusions he has drawn in Book III on the role of utilitarianism, for a key part of his overall argument for

⁴²⁵ The significance of Sidgwick's original intended neutrality, and the shift away from this neutrality that occurs as *ME* progresses, is fully addressed in chapter 4.

utilitarianism. Given that this very relationship is at the core of the central argument of this thesis, this chapter will end not with a discussion of the ultimate value that Sidgwick ascribed to utilitarianism (as chapter 2 did for deontology), but rather with the conclusion that Sidgwick's utilitarianism cannot be fully understood without reference to his deontological, or Intuitionist, principles. This approach is completely supported by Sidgwick and *ME*, in both structure and substance.

Overall, where deontological principles of morality were found in chapter 2 to be indispensable to Sidgwick's moral theory, but lacking in their ability to provide a complete system of ethics, this chapter ultimately finds utilitarian principles of morality to also be essential, but in need of a certain vital informing by deontological principles. Chapter 4 will then expound the particulars of this mutually informing relationship between deontological and utilitarian principles in full, as it appears within Sidgwick's theory of Rational Benevolence.

3.1. Utilitarianism in I:I : Right, Good, and Happiness

As was outlined in the chapter's introduction above, one of the most noticeable features of Sidgwick's utilitarianism as it appears at the beginning of *ME* is that it is often subsumed within discussions of other non-utilitarian notions and principles. This applies from the very beginning of the book, and most noticeably to Sidgwick's concepts of the cornerstones of absolute-principle based and end based ethical positions - 'right' and 'good' respectively. The preliminary outlines of these ideas provide a helpful miniature version of Sidgwick's later and deeper arguments regarding two of the most striking aspects of his version of utilitarianism – his view of rightness, and the identity of Happiness as a rational end. The following sections examine these early discussions, for their indications of the utilitarian line Sidgwick will take.

3.1.a. The Roles of 'Right' and 'Good' in I:I

It was noted in chapter 2 that early on in *ME*'s introduction, Sidgwick shows a preoccupation with right conduct (I:I:3). Here, the same passage is important for the converse reason of the apparently demoting effect that this has on the utilitarian concept of aiming at the good. Sidgwick states his position on the place of Ultimate

Good succinctly: No matter whether we take the view that good can be assumed of any ethical system, or the view (associated with the deontological position) that the performance of duty cannot be 'conditional on knowledge of its conduciveness to Happiness', 'the concept of Ultimate Good for man', Sidgwick says, 'would still not be important for the methodical determination of right conduct' (I:I:4).

There are two points that emerge here. The first, which is the more remarkable for a writer so widely considered to be a utilitarian, is that there is an early separation here between right action and the idea of *ends* – a separation in which ends do not necessarily determine the rightness of action, as traditional understandings of utilitarianism maintain⁴²⁶. Rightness appears here to be the superior concept for ethics, responsible as it is for establishing moral action, and utilitarianism's traditional method of relying on ends or the notion of the good to determine action thus appears to be side-lined very early⁴²⁷. The second point is that Sidgwick *does* hold good to be essential. At both the beginning and the end of these paragraphs, Sidgwick writes about the fundamental role of the good. It is, in his words, indispensable to the completeness of an ethical system (I:I:4).

This introductory outline of Sidgwick's understanding of the respective roles of right and good in ethics is actually a microcosm of his wider argument made throughout *ME* overall. Sidgwick's separation between rightness and goodness here points to his later emphasis (already discussed in chapter 2) on what is *right* being that which is objectively *rational*, and his subsequent use of his version of the Kantian maxim. His argument for the necessity of good in ethics is indicative of his later, more traditionally utilitarian, argument for Happiness. Yet despite this separation, 'right' and 'good' also appear to be in some kind of inextricable relationship. These two points together directly support my overall argument that

⁴²⁶ There is nothing immediate here that says that deontological/absolute-principle based systems must provide these rules for right conduct, as might be expected from the traditional association between such systems and the prioritising of 'right' - they could well be the rules of a comprehensive utilitarian system. Indeed, in a pre-*Methods* paper of 1873 paper simply entitled 'Utilitarianism' (this paper was previously unpublished until Marcus Singer's (ed.) 2000 volume *Essays on Ethics and Method* (Oxford: Clarendon Press: 2000) pp.3-9 – I will here after refer to this paper as 'Utilitarianism 1873'), Sidgwick states that he understands utilitarianism to 'supply a principle and method for determining the objective or material rightness of conduct'. However, Sidgwick's point at I:I:3 is that good *itself* is inadequate for determining such rules, and in making this point Sidgwick is still drawing a clear division between good, and right conduct.

⁴²⁷ In his well known paper, 'Some Fundamental Ethical Controversies' (*Mind*, Vol.14, No.56 (1889) 473-87), Sidgwick summarises this position with the denial that the idea of right can be 'explicable by the idea of good'.

both deontological and utilitarian ideas are essential to Sidgwick's moral theory, and that they are somehow inherently connected.

But, returning to these early pages of *ME*, the ambiguity in which Sidgwick's actual utilitarianism is often shrouded has already appeared. He initially points to the superior importance of good, but then waives it by drawing attention to the importance - practical and theoretical - of rules, and even of unconditional rules. He then reinforces the good again, drawing attention back to the significance of good's enduring presence, and the inclusion of Happiness as a vital part of its identity. Above all else, it is evident here that Sidgwick's views on those fundamental mainstays of utilitarianism – the notion of good and the idea of simple *means* to ends – are somewhat more flexible than might be expected of a standard utilitarian position.

3.1.b. Rightness in I:I: Happiness as a Rational Ultimate End

Continuing to outline the arguments that he will later make in full, Sidgwick's attention turns now to that question which he has already identified to be latent in any ethical system – that of *why* we should do something. The language of 'good' here disappears, replaced by that of 'reasonable', and 'reasonings' (I:I:5, 6). That we generally pursue courses of action that we deem to be *reasonable* is of no question for Sidgwick, and this connection between Reason and the aim of action appears even before he has specifically defined the utilitarian position as being that which aims at a certain end. Clearly, the concept of Reason appears fluid enough to Sidgwick to apply to the basis of any ethical system. This is of course an early indication of the shared appeal to Reason that will inform Sidgwick's own moral theory but for the present discussion which aims to demonstrate the uniqueness of Sidgwick's utilitarianism, it must be noted that the utilitarian principle just defined is only developed *from* Sidgwick's view on the role of Reason in moral processes.

This happens as follows. Sidgwick challenges the idea that our most common usage of 'ought' does not always suppose an ulterior end. In most cases, he states, this ought *is* actually a hypothetical imperative, and the 'end' (personal Happiness, in

Sidgwick's example) is merely optional (I:I:6)⁴²⁸. But this, he argues, is not a complete account of the notion of 'ought' – for we believe that it would be *irrational* not to care about Happiness, which thus makes the rational pursuit of Happiness an obligation. 'Ought' is now no longer relative; Happiness has become an ultimate end prescribed by reason, and as such Happiness is now a *categorical* imperative, 'prescribed...as Kant would say...without any tacit assumption of a still ulterior end' (I:I:7). That is, Happiness as a rational ultimate end has emerged as a non-consequentialist concept, in that it is *itself* right without reference to anything further⁴²⁹. From here, Sidgwick attaches that ought to the wider general Happiness, and thus arrives at his first formal definition of his own utilitarian position, which emphasises that general Happiness, rather than private, is this ultimate end categorically prescribed (I:I:8)⁴³⁰.

⁴²⁸ Happiness is actually one of two 'rational ultimate ends', the other being Perfection, or Excellence of human nature (I:I:9). This indicates Sidgwick's view that rational *ends* appear in deontological-duty or virtue based systems, as well as utilitarian ones, which is in itself an early suggestion of the common ground between deontological and utilitarian moral approaches in Sidgwick's thought. The idea of 'Happiness as a rational ultimate end' being somehow unique to Sidgwick requires clarification however. Obviously, Bentham himself held Happiness to be the rational ultimate end of being, and Sidgwick expressly points out that he is working from Bentham's argument that "the constantly proper end of action on the part of any individual...is his real greatest happiness..." (Bentham, *Memoirs* (in, John Bowring (ed.) *The Works of Jeremy Bentham Vol. 10* [1839] (London: W.Tait:1843) p.560, in *ME*, I:I:10). A separation between Sidgwick and Bentham must therefore be made, and this is done on two counts. The first is by drawing attention to the fact that Sidgwick draws this direct parallel between Happiness as an ultimate end and the 'ought' that is commonly associated with deontological ethics. This produces a far more robust defence of Happiness as an ultimate end than Bentham's. Secondly, Sidgwick's account of Happiness and the extent of its relationship with the more basic Benthamite notion of pleasure is more complex than Bentham's. This becomes clear as *ME* progresses, and is explored in full below, at 4.2.a.

⁴²⁹ This establishment of Happiness as a rational, non-consequentialist and ultimate end will be what informs the universal good inherent to Sidgwick's principle of Universal Benevolence, when he finally derives Universal Benevolence from his philosophical intuitions in Book III.

With the examples Sidgwick gives here – of there being a 'manifest obligation' to pursue one's own happiness – Sidgwick is clearly establishing the basis of Rational Egoism. This is not problematic however, as it is this argument for the reasonableness of Happiness as an ultimate end that Sidgwick applies directly to his argument for *general* Happiness, which is the basis of Rational Benevolence.

⁴³⁰ Sidgwick's whole definition of the general utilitarian position is as follows: "It is contested by many Utilitarians that all rules of conduct...are really – though in part unconsciously – prescribed as means to the general happiness of mankind...; and it is still more widely held by Utilitarian thinkers that such rules...are only valid so far as their observance is conducive to the general happiness" (I:I:8). In terms of his own view of utilitarianism, Sidgwick adds the qualification that "if the duty of aiming at the general Happiness is thus taken to include all other duties as subordinate applications of it, we seem to be again led to the notion of Happiness as an ultimate end categorically prescribed, - only it is now General Happiness and not the private happiness of any individual. And this is the view that I myself take of the Utilitarian principle" (I:I:8). Sidgwick expands on this in III:XIII, and I will return to his definition at the relevant point below. Here it is interesting to note Sidgwick's inclusion of the clause, 'though in part unconsciously' in his wider definition of utilitarianism. This seems to suggest that people could actually be following utilitarian principles, even when they feel themselves to be following more absolute, fixed principles, and that as such, all morality could eventually be found to be at least unconsciously utilitarian. Sidgwick himself considers this very possibility later in *ME*. I

The first time then that the traditionally utilitarian concept of Happiness appears in Sidgwick's work, it appears in inextricable relationship with the traditionally deontological categorical idea of 'ought'. Even more specifically, the *end* of Happiness has been established alongside the traditionally deontological view of principles, such as Veracity, which are also prescribed categorically (I:I:8)⁴³¹. Where Sidgwick's idea of Happiness is recognisably utilitarian then, the process through which the idea is established is not⁴³².

Sidgwick's references to Utilitarianism throughout the remainder of the introduction are rather more standard. He defines as Egoistic and Universalistic hedonism 'the two methods which take happiness as an ultimate end' (I:I:11), with the latter of these being that which is most commonly understood by the term 'Utilitarianism'⁴³³. He also closes I:I with a series of statements that re-emphasise utilitarianism as a vital aspect of morality⁴³⁴. But it is now possible to identify that one

acknowledge that this appears to provide strong support for the common belief that Sidgwick was ultimately a utilitarian. However, I also argue that even if the idea that Sidgwick believed all morality to be unconsciously utilitarian is maintained, the peculiar nature of that utilitarianism, given to it by its dependence on deontological principles, makes it possible to also state that Sidgwick believed all morality to be unconsciously deontological. The basis for this claim is made in chapter 4, where it is argued that neither utilitarian nor deontological principles are fully functional without crucial reliance upon the other.

⁴³¹ This is the direct result of Sidgwick's developing idea that the rational 'ought' applies just as much to any system of ethics.

⁴³² On the subject of ends, it must be noted that the idea of ends usually naturally involves the idea of *means* to those ends. This is, of course, a fundamental characteristic of the utilitarian position. However, the specific idea of means is conspicuously absent from I:I. This is prophetic of *ME* as a whole (see 4.2 below); Sidgwick is seldom obviously concerned with means, focusing almost entirely instead on the actual identity of the ends. There is admittedly, in Sidgwick's earlier words that in order to establish right conduct we must know the end, a latent sense of a utilitarian idea of means (I:I:3). This is actually the only place in which the phrase 'right means to its attainment' appears, and it does appear – as per traditional utilitarian understandings – ahead of 'right conduct'. But as was seen in that part of the discussion, Sidgwick waives the notion of good – which, if made superior, might legitimise any action provided it pertained to that good – in favour of the importance of what it is right to *do*. This is then followed of course with Sidgwick's view on what is meant by 'ought', which further elevates moral action away from being 'means', and towards a status in which the rightness of actions is not dictated simply by their conformity to ends. It is certainly not the case that Sidgwick adheres to a basic utilitarian standard of arguing that actions are only moral in as far as they are means to non-moral ends. Where Sidgwick *does* allow for the idea of means to ends is discussed shortly.

⁴³³ Sidgwick identifies Bentham with the origin of the doctrine for the first time here.

⁴³⁴ These pages were quoted in the introduction and in chapter 2 as evidence of the extremely similar starting points of Sidgwick and this thesis – that both deontological and utilitarian principles of morality appear to be relevant to day to day moral decision making, and that they often appear blended and sometimes confused – but they are worth quoting again here, for the sake of pointing out Sidgwick's belief that utilitarianism is a fundamental part of ordinary moral reasonings, alongside other ethical methods. 'The truth seems to be that most of the practical principles that have been seriously put forward are more or less satisfactory...so long as they have the field to themselves. They all find a response in our nature; their fundamental assumptions are all such as we are disposed to accept...When I am asked... "Have you not a moral sense?", "Do you not intuitively pronounce some actions to be right and others wrong?" "Do you not acknowledge the general happiness to be a paramount end?" I answer 'yes' to all these questions" (I:I:14). This passage also reinforces the sense

of the most noticeable characteristics of Sidgwick's particular utilitarianism is his emphasis on the role of Reason, and rationality. It is predominantly this that from the outset makes his position so distinctive compared to that of Bentham and Mill. Bound up with an often deontological, non-consequentialist view on the role of rationality in ethics, Sidgwick's adherence to this particular view is at the core of his two central pillars - his understanding of rightness, as will be seen, *and* his understanding of the general Happiness that is fundamental to utilitarianism. As *ME* progresses, Sidgwick deepens this rationality-based connection, that will serve to close the distance between utilitarian and deontological ideas.

The next section is concerned with where Sidgwick's view of Reason applies to his more detailed interpretation of rightness, the effect that this interpretation has on some of the key concepts in utilitarianism in general, and those concepts as they appear in Sidgwick's own utilitarianism.

3.2. The Rejection of Right as 'fitness to some end'

Only one short chapter of *ME* elapses before Sidgwick returns to the idea in I:III of what is right also being that which is considered to be reasonable, or rational. As was seen in chapter 2, Sidgwick first swiftly dispenses with the idea that 'the fundamental notion represented by the word 'ought' or 'right' can be discerned via any experience of the physical world (I:III:25). Then, following on from the absence of 'means' in I:I, Sidgwick also dismisses the idea that rightness can be an attribute of means (I:III:26). Again, this point was addressed in chapter 2 when the argument was being made that Sidgwick's is a deontological understanding of rightness, but here the passage is examined in terms of what it means for Sidgwick's utilitarianism. For if it is quoted in full, there appears a considerable challenge to a consequentialist viewpoint:

'It is urged that "rightness" is properly an attribute of means, not of ends: so that the attribution of it merely implies that the act judged right is the fittest or only fit means to the realisation of some end...and similarly that the affirmation that anything 'ought to be done' is always made with at least tacit reference to some ulterior end...but it seems clear (1) that

that there is a certain closeness between absolute-principle based and end based moral principles, as they appear in day to day morality.

certain kinds of action...are commonly held to be right unconditionally, without regard to ulterior results and (2) that we similarly regard as "right" the adoption of certain ends...In either of these cases the interpretation above seems clearly inadmissible' (I:III:26)

Sidgwick then makes his statement that 'we have therefore to find a meaning for 'right' or 'what ought to be' other than the notion of fitness to some ulterior end' (I:I:26).

Sidgwick's argument that rightness can be found in ends of course accords fully with the typical utilitarian position, held by both Bentham and Mill, that value lies in ends such as those Sidgwick references here. But the difference between Sidgwick's thought and that of the utilitarianism typical of his time (and of most 20th Century utilitarianism) lies in what Sidgwick's definition of rightness *excludes*. Sidgwick's mention of the 'adoption' of these ends suggests he recognises that means to those ends must have a role somewhere- but his uncompromising statement that rightness cannot mean 'fit to an ulterior end' appears here to preclude the idea that actions can have the property of 'rightness', merely provided they are pertaining to a certain end⁴³⁵. When it is considered then that Sidgwick is held to represent the epitome of classical utilitarianism, this position on 'rightness' as not applicable to 'fit to an end' is striking. Sidgwick may be consistent with classical utilitarianism in his recognition of ends, but he is apparently *not* consistent in terms of how he sees the relationship between action and those ends. It is only the ends themselves that qualify as 'right' in this unconditional and categorical sense (which is the same 'right' that is used to describe certain unconditional and non-consequentialist duties): Means, according to Sidgwick at this stage, cannot qualify in the same way. Ultimately, according to Sidgwick, 'ought' and 'right' cannot actually be fully defined in express terms at all (I:III:32), let alone be used to describe relative or conditional actions. All this results in an extremely unusual position for a utilitarian⁴³⁶.

⁴³⁵ Sidgwick is as good as explicit on the exclusion from his understanding of 'rightness' as involving means to ends. By including the word 'similarly' in his reference to how we regard the adoption of certain ends, Sidgwick draws an unequivocal parallel between certain ends – which so commonly invoke the notion of means in the utilitarian tradition - and the recognisably deontological view that certain actions are right 'without regard to ulterior results' (I:I:26). That is, Sidgwick means that these certain ends are also considered to be right *unconditionally* and *in themselves*.

⁴³⁶ Sidgwick gives an even more direct defence of his idea of rightness as not simply meaning fit to an ulterior end in 'Fundamental Ethical Controversies'. Here, Sidgwick writes that a "rational judgement" is normally expressed in the form 'x is right', or 'x ought to be done'; and if the judgement be attained by deduction from a principle, such a principle is always capable of being expressed as a proposition in which the word 'right' or 'ought' occurs' (p.482). He continues, in discourse with Fowler, that he believes the notions expressed by 'ought' and 'right' are 'ultimate and unanalysable', and again here explicitly rejects the idea, this time put forward by Fowler, that right could mean 'fit to the realisation of

There are further consequences of this view of rightness, for three more of traditional utilitarianisms' most basic components: The principle of utility, the consequentialism inherent in that principle, and – overall – utilitarianism's teleological nature. The challenge of Sidgwick's work to these cornerstones of utilitarianism is addressed in the following sections.

3.2.a. The Rejection of Right as 'fitness to some end': The Challenge to the Principle of Utility

Sidgwick's exclusion of means from his definition of rightness immediately calls to attention a foundational tenet of the classical utilitarianism of Sidgwick's time - Bentham's principle of utility. This is 'that principle which approves or disapproves of every action according to the tendency which it appears to have to augment or diminish the happiness of an individual...' ⁴³⁷. In other words, actions are to be judged morally right or wrong in as far as they are the means to either securing Happiness, or avoiding pain.

Compare the principle of utility to Sidgwick's words, 'the attribution of [rightness] merely implies that the act judged right is the fittest or only fit means to the realisation of some end...' (I:III:26), and his immediate denial that this emphasis on means to ends provides the correct version of rightness ('...the interpretation above suggested seems clearly inadmissible'). The disagreement between the two positions is clear. The principle of utility is plainly queried by Sidgwick's defining right *not* as fit to an end, and by his rejection of the idea that rightness is an attribute of means ⁴³⁸. If rightness has this absolute, unconditional quality, then actions cannot

some end'. Sidgwick is at his most emphatic when he writes, 'the approval of the conscious choice of another's greater good in preference to the chooser's lesser good, is regarded [by Fowler] as a normal moral judgement...this judgement must be expressible in the proposition that 'conscious choice &c. is right', and the word 'right' in this proposition cannot *mean* 'conducive to the greatest good on the whole'. On this basis, Sidgwick concludes, the idea of right cannot 'be explicable by the idea of good' (p.485). Lastly, and with particular significance for this argument which aims to show Sidgwick's utilitarianism to be unconventional, Sidgwick notices that on this point about rightness, other utilitarian writers do not agree with him.

⁴³⁷ Bentham, *An Introduction to the Principles of Morals and Legislation* (1789) Ch.1, §4 (in, John Bowring (ed.) *The Works of Jeremy Bentham Vol. 1* [1839] (London: W.Tait:1843)

⁴³⁸ Sidgwick recognises the difference between his thought and Bentham's here, and in a footnote to his definition of rightness draws a comparison between the two positions himself. Sidgwick's point in the note is that Bentham cannot mean that rightness is simply 'conducive to general happiness', as this would 'hardly serve as the fundamental principle of a moral system' (I:III:26). Sidgwick departs significantly from Bentham then, over Bentham's definition (or lack, thereof) of the *kind* of rightness that Sidgwick argues for in I:III – for it is clear from Bentham's writing that this *is* what he meant.

simply be called right when they pertain to a certain end, and the fundamental precept of utility is refuted.

In this way, the consequentialism inherent in the principle of utility is also contested⁴³⁹. Consequentialism holds that an action can be judged to be right when it brings about certain desirable outcomes. It has already been argued in chapter 2 that as Sidgwick argues for the distinct nature of rational, ultimate *ends* as being *right* without regard to ulterior results, an ultimate rational end such as Happiness is actually in itself non-consequentialist - and here non-consequentialism *also* seems to apply to actions⁴⁴⁰. Actions that bring about certain outcomes are simply fit to an end, and this does not accord with Sidgwick's definition of rightness.

Bentham says, 'Of an action conformable to the principle of utility' (which approves or disapproves actions so far as they augment or diminish happiness), 'one may always say either that it is one that ought to be done, or at least that it is not one that ought not to be done. One may say also, that it is right it should be done...thus interpreted, the words *ought*, *right*, and *wrong*, and others of that stamp, have meaning: when otherwise, they have none'. Bentham does therefore, as Sidgwick suggests, hold that by 'right' it is meant 'conducive to the general happiness'. As Ronald D. Milo points out in his 1974 article 'Bentham's Principle' (*Ethics*, Vol.84, No.2 (1979) 128-139) it appears here as though Bentham is not just proposing a normative test for rightness, but that he is also making a meta-ethical statement about what it means to say that an action is right or wrong. (Milo also points out Sidgwick's dispute with Bentham on this point (p.132). On this comparison then, the disagreement between Bentham and Sidgwick over what is meant by 'right', and an unavoidable subsequent disagreement over the accuracy of the principle of utility, becomes clear.

Sidgwick specifically discusses the principle of utility – and Mill's defence of it - later in *ME* (III:XIII:387), and does actually offer his own interpretation of the principle. This appears deep within the very particular context of Sidgwick's argument for the *proof* of the utilitarian principle. See chapter 4.

⁴³⁹ Given that consequentialism holds that the morally right action is the one that brings about the best overall outcome, and that the rightness of actions is determined by their consequences, consequentialism is inherent in the principle of utility.

It should be noted that consequentialism, as a theory of ethics, is not completely limited to utilitarianism (see chapter 1 p.94 of this thesis' for G.E.M. Anscombe's first use of the term). The basic premise is that the consequences of one's actions are the ultimate standard for judging rightness and wrongness – but consequentialist theories in general do not necessarily contain specific definitions of what those consequences are to be. As such, theories such as ethical egoism, in which actions may either benefit, disadvantage or be neutral towards others, are not excluded from the equation. (Sidgwick's own exploration of ethical egoism concludes that the consequences of egoistic conduct may at times promote the general welfare, but only by proxy). Nor does it exclude ethical altruism, in which the consequences of an agent's actions pertain to everyone except that agent. For the present purposes however, the position is taken that utilitarianism is generally considered to be the paradigmatic example of a consequentialist ethical theory.

⁴⁴⁰ It is also noticeable that specific references to 'consequences' are remarkably sparse in *ME*, appearing at only three places in Book I, and often in conjunction with simply defining utilitarianism as opposed to Intuitionism. 'Consequences' appears at I:I:8, where utilitarianism is being defined negatively in relation to deontology; I:III:35, where Sidgwick argues that no-one really means to deny that we are obliged to certain actions without *any* regard to consequences; and I:VI:87, which is the only place at which Sidgwick uses the direct phrase 'utilitarian consequences'. One way in which this might be explained is that arguably, the notion of consequences is always implied in Sidgwick's references to actions that pertain to ulterior ends. But Sidgwick's apparent neglect of this essential part of the utilitarian doctrine is perhaps better accounted for in another way, that is more telling of his moral theory overall. This is the argument that Sidgwick is purposefully maintaining a slight distance between what I will call the 'pure' consequentialist idea that consequences – 'ends' - are all that

It is true that Sidgwick has identified certain *ends* as being the ultimate reasons for acting, and that this could be seen as indirectly maintaining the underlying idea that ends must dictate action to some extent. This is explored fully below. But within the context of this current discussion which seeks to demonstrate that Sidgwick's utilitarianism is not typical of his time, the classical tradition, or indeed of much utilitarianism that has followed, there is strong evidence here for an apparent divergence between Sidgwick's definition of rightness, and the cornerstone of the traditional utilitarian doctrine. Phrased in the most extreme terms, it would seem as though the role of the principle of utility in determining moral action is, for Sidgwick, seriously reduced.

3.2.b. The Rejection of Right' as 'fitness to some end': Sidgwick's Teleology

If the principle of utility is thus contested, then there emerges a related but wider consequence for Sidgwick's utilitarianism. This is that the entire teleological nature of utilitarianism as understood by Sidgwick could be called into question.

Teleological systems of ethics, of which utilitarianism is the foremost example, maintain that moral rightness lies in the bringing about of certain ends. If actions that pertain to ends cannot be called 'right', then a fundamental part of teleology is refuted. This argument does, however, depend on what is meant by 'teleological' in ethics. If it is taken to mean that the moral rightness of an act is determined by its being fit to some end, then Sidgwick's definition of rightness does disqualify his model from being teleological⁴⁴¹. This view that the teleology in Sidgwick's theory is brought into question by his understanding of rightness is further supported by his Reason-based process of establishing ultimate rational ends, and by the *type* of value that he places on those ends. This is because Sidgwick's argument for what is reasonable being also that which is right shares a striking amount with Kant's prioritising of the right. Even if no connection between Sidgwick and Kant is actually

matter in *any* situation, and his own emphasis on Happiness as the only true *right* rational end. This further contributes to the impression already developing that Sidgwick is not – and will not be – building a version of utilitarianism that simply revolves around the importance of consequences, or the bringing about of certain results.

⁴⁴¹ This version is based on understanding ethical teleology in its rawest form, which can be represented by the phrase 'the ends justify the means'. I do not believe that Sidgwick's concept of rightness allows for such an understanding, and that Sidgwick therefore avoids this particularly difficult area of utilitarian ethics.

made, it still remains that Sidgwick attributes the same kind and degree of rightness to Happiness as he recognises to be given to certain moral duties by the Intuitional (deontological) position, and in both these cases rightness is wholly independent, and cannot be achieved or established by anything conditional. To this understanding of teleology in ethics then, Sidgwick does not seem to conform.

Alternatively, teleology can also be construed as holding that intrinsic value lies in (usually non-moral) ends, and that moral obligation can then be *derived* from these ends. This understanding places the emphasis not on the actions, but on the moral status of the ends themselves. In this case, Sidgwick's emphasis on Happiness as an ultimate end that is right in itself and rational to pursue does clearly accord with a teleological understanding. It is a stricter teleology than the first definition, and it is the only one that Sidgwick's definition of rightness allows - but it is also the route through which 'rightness' can be extended, on a very specific understanding, to actions that pertain to an end⁴⁴². Sidgwick explains this via his discussion of Truth-speaking. Here, he allows that means such as Truth-speaking do acquire a certain 'ought', but *only* by extension of the reasonableness of a rational end (I:III:35-36). To expand on this, Sidgwick's justification of the *means* of Truth-speaking can only happen on the grounds that the end (of preserving society) is first recognised as a rational end; as it is only rational for a rational being to aim at rational ends, these sorts of means take on the same sort of 'ought' (i.e. moral obligation) outlined in I:I⁴⁴³. Means are therefore restored, and Sidgwick's utilitarianism emerges more recognisably here, as does a teleological approach⁴⁴⁴. But the ends must be

⁴⁴² This is Sidgwick's reintroduction of means to ends into his definition of rightness, that was indicated by chapter 2.

⁴⁴³ This is the 'ought' which is inherently connected to Reason. This 'ought', applied now to those means, *still* brings with it the properties peculiar to that notion, of being a dictate of reason, and as being something which is only rational for a rational being to pursue. (Sidgwick's keenness to convey the fact that the dictate of Reason applies just as much to means as to ends appears plainly, and with a note of frustration: "and the notion of 'ought' as used in in either dictate is that which I have been trying to make clear"). Sidgwick continues this point on p.35, where it might be argued that his phrase 'the recognition of an end as ultimately reasonable involves the recognition of an obligation to do such acts as most conduce to the end' is a clear defence of the utilitarian idea that actions are right in as far as they conduce to a certain end. This also seems to be supported by a rather throwaway comment from Sidgwick made in the preceding paragraph, in which he states that 'actions are 'reasonable' or (in an absolute sense), 'right''. Both of these passages appear to validate the idea that Sidgwick believes any reasonable action to be 'right', because it is reasonable. But in response to this I argue that even this understanding is dependent on the prior establishment of ultimate, rational ends.

⁴⁴⁴ It is certainly teleological in contrast to pure deontological theories in which an action is held to be right (or wrong) regardless of *any* ulterior results. But the distance between the two positions might be reduced again with Sidgwick's comment that there has never been a type of morality that did not

determined first, as Sidgwick's apparent prioritising of the right outlined above also demands - and these ends must possess the specific identity of being ultimate, rational ends. The importance of the distance that Sidgwick has maintained between his own concept of ultimate, rational ends on the one hand, and the more traditionally utilitarian concept of consequences on the other, now becomes evident. Consequences and rational, ultimate ends do not have the same moral value in Sidgwick's thought; only ends that have been established *rationally* have that peculiar and absolute non-consequentialist rightness⁴⁴⁵. The priority of the right appears clearly here again, with the usually utilitarian idea of means to ends only emerging from a traditionally deontological and non-consequentialist understanding of exactly what is meant by the term 'right'. Means still depend entirely on the prior recognition of an ultimate, reasonable end. But if rational ends are established correctly, means to those ends also acquire an 'ought', which denotes rightness⁴⁴⁶.

Overall then, Sidgwick's theory of moral obligation is teleological in that obligations can be derived from intrinsically valuable ends. On the other hand, it is not teleological in that moral rightness itself cannot be determined *purely* by what is fit to an end. It is my view that Sidgwick's theory can be described as either teleological or non-teleological – and this in itself is not typical of a classically utilitarian position.

consider consequences to *some* extent (I:VIII:96). On this basis, Sidgwick's argument for rational ends is purely a logical result of that observation, and could apply equally to absolute-principle based systems – which must also include ends to *some* degree.

⁴⁴⁵ Strictly speaking, Sidgwick does point out that the obligation is not quite unconditional in the same way as duties are considered to be unconditional by the Intuitionist position. But, he says, the obligation in the case of what is conducive to an end does not depend on any non-rational desires/aversions (I:III:35). It therefore still falls within the remit that Sidgwick describes at I:III:25 (and which he alludes to earlier at I:I:9) of what is meant by the 'fundamental notion' of 'ought', which has a unique character, excludes any reference to sensory/psychical experience, and therefore has that 'peculiar significance' of rightness.

⁴⁴⁶ Via the same route, a consequentialist aspect can also be returned to Sidgwick's theory. As consequentialism holds that the consequences of an action are the basis of judgement for the rightness of that action, then it can be taken that actions that are imbued with an 'ought' *because* they are aiming at an ultimately reasonable end are right according to a consequentialist viewpoint. When this argument is expanded, it can be seen that Sidgwick's only allowing 'consequences' to be understood in this way prevents a situation in which *any* end *perceived as positive* would justify an action. The most obvious example with which to demonstrate this point is that of Nazi Germany. There, the idea was conceived and encouraged that mass genocide would be beneficial to the country; those beneficial consequences were used as the justification for the actions of the Third Reich, when it is clear that such actions do not conform to any objective, rational standard of rightness. On Sidgwick's method of establishing the rational ultimate ends that are to dictate action, the broad consequentialist concept that ends simply justify the means is, as Sidgwick would say, clearly inadmissible. This limitation to which ends can be considered rational and therefore morally right is also relevant to the possible establishing of limitation to action in the context of the Parental Predicament, a point that I will explore in the thesis conclusion.

3.2.c. Conclusions on Sidgwick's Definition of Rightness

It was seen in the preceding section that despite Sidgwick's early denial that right means 'fit to an ulterior end', Sidgwick's view of rightness, and how the moral obligation established by that view applies to ends, can in some respects be described as teleological. But this question regarding the apparently debatable status of Sidgwick's teleology that is raised by his definition of the right seems to be a valid one – and, crucially, the responses to it outlined here are important for what they reveal about Sidgwick's emerging utilitarianism. This importance can be summarised with the statement that all of the predominantly utilitarian and teleological concepts discussed in this section only appear in Sidgwick's understanding in highly specific - and restricted - ways. That is, it is only on a particular understanding of *certain specifically defined ends* that are ultimately rational i.e. established as right without reference to ulterior results, that means acquire the 'ought' that denotes rightness at all. Sidgwick's teleology is therefore highly dependent on his prior view of the right, and his thought here treads a fine and sophisticated line between teleological utilitarianism, and a non-consequentialist priority of the right.

As for the principle of utility, Sidgwick's teleology is not as unmistakeable as Bentham's, or Mill's - and consequently neither is his use of the principle of utility. Whereas reasonable actions that pertain to a reasonable end are found to have some kind of moral worth (and although Sidgwick agrees with Bentham that Happiness is the rational end in question) the principle of utility simply does not seem adequate to capture the precise origin and degree of this worth. It is at best an incomplete theory as to how to determine the rightness of actions. In any case, Sidgwick certainly offers a more refined account of that principle's role than either of his predecessors.

This leads to what can be said more positively of the effect that Sidgwick's position on rightness has on his utilitarianism overall. It is still Sidgwick's view of rightness that, it might be better to say, *cannot simply* be defined as fit to an ulterior end that makes his writing here unique. He rejects the standard utilitarian idea that rightness is simply fit to an ulterior end, but his particular view of rightness, which relies on what can be said about the notion of 'ought' eventually allows for a system

in which both ultimate rational ends, and certain related rational actions, are right⁴⁴⁷. Phrased in terms that pertain to my argument in this thesis that Sidgwick's moral theory is comprised of both deontological and utilitarian principles, Sidgwick has perceived rightness in both approaches, and connected them in an intricate balance which is maintained by the precedence of that unconditional and categorical rightness. And in the narrower terms of Sidgwick's utilitarianism - if it can be called that at this stage - this position makes it more robust, better supported, and ultimately perhaps more morally and philosophically credible than that of his time.

3.3. Happiness and Pleasure as Ultimate Ends

The last issue raised by Sidgwick's definition of rightness, and his highly specific position on the rightness of means to ends, is that of Sidgwick's views on quite what is meant by Happiness as an ultimate end. Specifically, Sidgwick considers how far *pleasure* is to be included in this definition. On this topic, Sidgwick reaches the first of the most obvious utilitarian concepts to have appeared in *ME*⁴⁴⁸. But here, again, Sidgwick is not entirely in line with the traditional utilitarian understandings of the classical era. Firstly, and contra Bentham, Sidgwick rejects pleasure as an ultimate end. Secondly, he offers a defence of Happiness as an ultimate end that supersedes those of both Bentham and Mill- while also bringing him back into line, rather surprisingly, with Bentham. These points are explored in turn, in order to ascertain Sidgwick's view on this fundamental utilitarian idea, and the place that he affords it in his own version of the doctrine. This will be crucial for the eventual expounding in chapter 4 of Sidgwick's moral theory of Rational Benevolence which depends on

⁴⁴⁷ As Sidgwick himself says at I:III:32-33, the notion which 'right' and 'ought' have in common is too elementary to admit of any formal definition. It follows, therefore, that Sidgwick should find rightness to apply in both deontological and utilitarian viewpoints, as he has done here. Furthermore, Sidgwick himself states in 'Fundamental Ethical Controversies' that what words such as 'right' and 'ought' have in common is the same in different ethical systems... That is, Sidgwick does expressly state his recognition that this understanding of rightness transcends the traditional divide between deontological/utilitarian understandings of right.

⁴⁴⁸ Until this point in *ME*, utilitarianism has been held in relatively low profile, with concepts primarily considered to be deontological often dominating the discussion, and with any utilitarian ideas mainly being subsumed therein. Where there has been deeper and prolonged discussion of Sidgwick's utilitarianism, as throughout 4.2., I have raised those questions myself, in order to state what can be inferred about his utilitarianism at this stage.

Sidgwick's understanding of Happiness as the good inherent in Universal Benevolence.

3.3.a. Sidgwick's Rejection of Pleasure

It has already been noted above that Sidgwick considers Happiness to be one of only two rational ultimate ends (see I:I:9, 11). On this, he is emphatically clear. What is often less clear (at least until a complete reading of *ME* has been made, as argued below) is exactly how Sidgwick sees the relationship between the concepts of 'Happiness', which is considered to be a rational ultimate end, 'hedonism', which is used by Sidgwick to denote those methods that hold Happiness to be an ultimate end, and 'pleasure', the idea that accompanies hedonism. For the term 'hedonism' itself represents the theory that pleasure is intrinsically valuable (and that pain is intrinsically non-valuable), and this is the position of Bentham, who famously asserted that 'nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*', and that 'it is for them alone to point out what we *ought* to do, as well as determine what we *shall* do'⁴⁴⁹. Bentham then equates this pleasure directly with Happiness. In Bentham's doctrine then, Happiness itself is quite clearly the hedonistic understanding of that word – and Sidgwick appears to follow Bentham's lead here in equating the two⁴⁵⁰. It would appear therefore as though Sidgwick is conforming to this idea that *pleasure* is synonymous with the Happiness that is the ultimate end. To some extent, this *is* Sidgwick's position. On the other hand, he expressly denies that pleasure is an ultimate end. In order to understand this complication, and its significance for Sidgwick's utilitarianism, it is necessary to begin with the denial of pleasure as an ultimate end⁴⁵¹.

Sidgwick challenges Bentham's view that pleasure and pain are the sole motivators to voluntary action by pointing out that we frequently act on impulses that

⁴⁴⁹ Bentham, *An Introduction*, I:I – italics mine.

⁴⁵⁰ At I:VI:41, where Sidgwick embarks on the first discussion of pleasure and desire, Sidgwick specifically refers the reader back to his statement in *ME*'s introduction that he is using the 'exact hedonistic interpretation of 'happiness' which Bentham has made current', and which Sidgwick considers to be the 'most suitable use of the term' (I:IV:41).

⁴⁵¹ As mentioned in the literature review, there is some debate over whether or not Sidgwick believed pleasure to be the end of rational action. Thomas Christiano, for example, accepts without question that this is Sidgwick's position; Roger Crisp expressly denies it. I am in agreement with Crisp; pleasure is an integral part of Happiness, which *is* the end of rational action, but it is not the end itself, as I will argue in this section.

are not directly related to pleasure (or relief from pain) – *and*, conversely, that a very part of the pleasure that is in question is actually often brought about by acting on those non-pleasure related impulses (I:IV:47-48, 48-49)⁴⁵². Sidgwick concludes that ‘the doctrine that pleasure (or the absence of pain) is the end of all human action can neither be supported by the results of introspection, nor by the results of external observation’ (I:IV:53), and that there is ‘no necessary connection between the...proposition that pleasure or absence of pain...is always the actual ultimate end of my action, and the ethical proposition that my own greatest happiness or pleasure is for me the *right* [Sidgwick’s italics] ultimate end’ (I:IV:41)⁴⁵³.

Sidgwick’s claim that pleasure cannot be an ultimate end is not strictly original – J.S. Mill had already replaced the language of ‘pleasure’ as the end of human action with that of ‘happiness’ in his direct continuation (and defence) of Bentham’s work. But Sidgwick’s handling of ‘pleasure’ and ‘happiness’ arguably supersedes Mill’s. It is, firstly, a perfect example of Sidgwick’s avoiding of Moore’s ‘naturalistic fallacy’ which Bentham was thought to have committed by claiming that because people *do* desire pleasure, people *ought* to desire pleasure⁴⁵⁴. But furthermore, it is Sidgwick’s interpretation of rightness that allows a far more convincing explanation as to why pleasure cannot be the end of human action. Given that ‘rightness’ is derived from what is rational, ‘rightness’ is restricted to ends that can be determined via Reason, or rationality, only, and thus excludes ends that are derived from sensory experience. As Sidgwick himself points out, this necessarily excludes the idea of [the

⁴⁵² For example, we must avoid over indulging in the pleasure of eating in order to avoid ill-health which will not be conducive to real long term happiness, or, conversely, we must pursue certain less pleasant activities, such as exercise, for the same reasons. Sidgwick refers to this situation as the ‘paradox of hedonism’ (I:IV:48, 51, 52). Sidgwick also extends his challenge to pleasure as the sole motivator via his observations on sacrifice – namely, that an individual may well recognise that a sacrifice will not benefit themselves, and yet still determine that it must be made (I:IV:52). This is a real sacrifice, and not a simply exchange of short-term pleasures for long-term ones, as Bentham might have interpreted a sacrifice situation.

⁴⁵³ Even if it can be proven that each individual is led to do what, on his view, would be the most conducive to his own happiness, this still does not admit of the quality of ‘rightness’: The individual may indeed believe that they are acting reasonably, but they are acting from a psychological law, which of course cannot be given the status of being a dictate of Reason (I:IV:41). Sidgwick’s view that pleasure is clearly not always the ultimate end, and certainly not the end of right conduct, also appears in his paper, ‘Pleasure and Desire’, (*The Contemporary Review*, Vol. 19 (1872), 662-72).

⁴⁵⁴ Mill had famously attempted to overcome the problem of the naturalistic fallacy, but – at least according to Moore – very little success. Although Mill has since been largely exonerated of this charge, Moore’s influence has lasted in the impression that Sidgwick avoided the naturalistic fallacy with considerable more security than Mill. In fact, Moore was ready to claim that by refuting the idea that goodness is not definable or analysable, Sidgwick may actually have discovered the fallacy. See Moore, *Principia Ethica*, pp.17, 21. See also Thomas Baldwin, *G.E. Moore; The Arguments of the Philosophers* (Oxford: Routledge:1992) p.53)

experience of] pleasure as being, in the moral sense, 'right'. Sidgwick's position on pleasure stands in stark contrast then to Bentham's account that it is pleasure and pain that are to determine what we normatively *ought* to do.

The complication with this rejection begins when it is seen that Sidgwick, in that chapter which rejects pleasure as an ultimate end, states that he is using the 'exact hedonistic interpretation of 'happiness' which Bentham has made current' (I:IV:41), as he also referenced at I:I:10. It seems that this is done in order to remind the reader of the central concepts of the classical, Benthamite utilitarianism to which Sidgwick has regularly referred, and with which Sidgwick is working in this exposition of the various 'methods of ethics'⁴⁵⁵. But re-examining I:I with this possibility in mind only exposes the fact that Sidgwick never actually gives Bentham's interpretation of Happiness *per se* – he only quotes Bentham's view that Happiness is to be 'the true standard of right and wrong, in the field of morals' (I:I:10). It can only be surmised then that this 'exact hedonistic interpretation' to which Sidgwick refers is Bentham's understanding of Happiness. The major problem with this is that Bentham's definition of Happiness as an ultimate end *is* the pursuit or the attainment of pleasure, and the avoidance of pain. But with Sidgwick having specifically denied that pleasure/avoidance of pain can be an ultimate end, the Benthamite version of Happiness defined as pleasure cannot, by logical extension, be an ultimate end.

There is, then, a noticeable contradiction here between Sidgwick's and Bentham's respective interpretations of Happiness, that is then exacerbated by Sidgwick's claim that he himself is using Bentham's interpretation of 'hedonism'. So how is this rejection of pleasure, which seems to accord so well with Sidgwick's other points made so far, to be reconciled with his use of the Benthamite version of hedonism? The answer to this points to a true strength of Sidgwick's utilitarianism – and reveals the interpretation of Happiness that is going to underpin his own moral theory.

⁴⁵⁵ In another footnote on p.87, Sidgwick reiterates further that he has called the 'ethical doctrine that takes universal happiness as the ultimate end and standard of right conduct by the name of Bentham', because thinkers who have taught the doctrine have most frequently referred to Bentham as their master (I:VI:87). Of course, we can agree with Sidgwick that universalistic hedonism is indeed Bentham's doctrine (even if Sidgwick's interpretation of that doctrine is far more sophisticated and complex than Bentham's), but this does not remove the problem – at this stage of *ME* at least - of how that hedonism is to be defined.

3.3.b. Sidgwick's Interpretation of Happiness as an Ultimate End

In answering the question just posed, it must first be noted that throughout I:IV, Sidgwick uses 'happiness' and 'pleasure' interchangeably (see I:IV:41, 44, 51). This implies that Sidgwick does not seem to think that there needs to be a distinction between his own version of Happiness and Bentham's, and further compounds the apparent confusion noted above⁴⁵⁶. There is a degree to which 'pleasure' could be interpreted as being subsumed within the wider and greater notion of 'ultimate Happiness' as the individual actions that lead to Happiness, and this would mean that Sidgwick is really just using a Benthamite interpretation of Happiness after all. But that is not what is at stake at this particular point in the text: Sidgwick is specifically discussing the status of *pleasure as an ultimate end*, and he is rejecting this idea.

The theory that pleasure is included within the ultimate notion of Happiness is, however, the key to understanding the apparent contradiction, and is more effective if a wider view of *ME* is taken. To this end, we must accept the last definition of 'utilitarianism' that Sidgwick gives in I:I: "The two methods which take happiness as an ultimate end it will be convenient to distinguish as Egoistic and Universalistic Hedonism: and it is the latter of these, as taught by Bentham...that is more generally understood under the term 'Utilitarianism' (I:I:11). The use of 'hedonism' here is still awkward, in that it invites ideas of *pleasure* seeking etc. by way of definition. But whereas the language of hedonism is often accompanied throughout *ME* by the ideas of 'pleasure' and 'pain', it is *always* accompanied by the language of *Happiness* (see I:VI:84-85, II:I:120, III:XIV:402, IV:I: 413), which Sidgwick convincingly defines as a rational, ultimate end⁴⁵⁷. As *ME* progresses therefore,

⁴⁵⁶ In only two places does he use the terms 'pleasure' and 'happiness' separately, and even this separation is not made in express terms (see I:IV:51,52).

⁴⁵⁷ See III:VIII:380-384 (and also III:IV:241) for Sidgwick's specific defence of Happiness as the true rational end. There are also a few sections of *ME* that can add further weight to the argument being made here that it is Happiness and not pleasure that is the ultimate end. Firstly, pleasure is not mentioned once in *ME*'s introduction, where Sidgwick generally does include all of the notions most important to both the Intuitionist and the utilitarian points of view. 'Happiness', on the other hand, appears 17 times in the introduction in the context of being a rational ultimate end, as well as many other times as a general word. Even in these 'general' uses, Sidgwick never associatively uses the word 'pleasure'. Secondly, Sidgwick's definition of right can be recalled: Happiness qualifies under that definition because it is rational, and not fit to a still ulterior end, whereas pleasure is disqualified from the status of rightness on both of those grounds. Thirdly, at I:IX:107, Sidgwick rejects that 'good' is equivalent to pleasure, and by a process of elimination arrives at the conclusion that the good of man is Happiness (when it is not Perfection or Excellence of Human Existence) – not pleasure. Then

'hedonism' gradually takes on a character in Sidgwick's work in which pleasure and the avoidance of pain are integral parts of Happiness itself, but are *not* considered to be rational ultimate ends. The term 'hedonism' still represents Happiness, and by extension represents the pleasure that is essential to that Happiness. But, crucially, hedonism no longer refers directly to the pleasure *itself*. This is what distinguishes it from Bentham's use of the word. Bentham's hedonism holds pleasure as an *ultimate* end, where Sidgwick's does not. This explains why Sidgwick can both refer to Bentham's definition of Happiness, as he does in the note on p.87, *and* reject Bentham's view of pleasure. Sidgwick accepts Bentham's position that 'Happiness' is a rational ultimate end, and even that that Happiness involves hedonistic elements in its definition. But Sidgwick can differentiate that pleasure has a role/status wherein it has instrumental or non-intrinsic value only⁴⁵⁸. Sidgwick thus retains a form of hedonism in which pleasure and Happiness are intrinsically linked, but not strictly interchangeable⁴⁵⁹.

of course there is the passage already cited from I:IV, in which Sidgwick renounces the idea that pleasure is not always what amounts to Happiness (see also I:IV:48) As for Sidgwick's view on the true identity and role of pleasure/pain, the protracted discussion of this appears in Book II, which addresses egoistic hedonism.

The details of Sidgwick's view of pleasure per se need not concern us here, but it is worth noting that – somewhat surprisingly – Sidgwick agrees with Bentham and disagrees with Mill that pleasures are to be measured quantitatively, not qualitatively, as per Mill's attempt to refine this coarser aspect of Bentham's doctrine (see II:I:120-121, II:II:123-125. See also 'Utilitarianism 1873', where Sidgwick makes the same argument). This version of commensurability, Sidgwick argues, is necessary if Happiness is to be understood as the greatest surplus of pleasure over pain. Because this definition presupposes quantity as the standard of measurement, more pleasure must be sought as opposed to less regardless of any other qualities that those pleasures may have. His other reason for agreeing with Bentham is that for all pleasures to be understood as pleasures in the first place, they must have a common property, and that it is therefore possible to weigh amounts of this property. To say that something can be 'the greatest amount', while denying that the components can be commensurable would be, Sidgwick says, 'a mathematical absurdity' (IV:I:413).

⁴⁵⁸ In addition to his argument in I:IV that pleasure cannot be an ultimate end, and furthering the point made above that Sidgwick's definition of rightness as not fit an ulterior end denies pleasure the status of an *ultimate* end, on p.92 Sidgwick also states that he is rejecting Aristotle's definition of Happiness as 'eudemonia', given that this only really translates as 'well-being', and well-being could still be interpreted to mean a variety of things i.e. it is clearly not an ultimate end. Using these same grounds, it is possible to create a greater distance between Happiness and pleasure in Sidgwick's work: Pleasure, like well-being, can still be variably interpreted – it is therefore clearly subordinate to some greater and higher principle, which is that of Happiness.

⁴⁵⁹ Sidgwick does continue to occasionally use either Happiness or pleasure to refer to the same thing (see for example I:VI:78, II:I:120, IV:I:413) and at one point states that "Happiness (= sum of pleasures)" (III:XIV:407). This is also clear in 'Utilitarianism 1873', where Sidgwick states that "happiness" must be understood as equivalent to 'pleasure', and that 'Pleasure cannot be distinguished from Happiness'. I believe that this apparent equivalence can be explained however when it is borne in mind that pleasure remains the medium through which Happiness is actually understood. Sidgwick continues the passage just quoted from 'Utilitarianism 1873' with the qualification '...except that Happiness is rather used to denote a sum or series of those transitory feelings each of which we call a Pleasure. The utilitarian, then, aims at making the sum of preferable or desirable feelings in the world...as great as possible'. Sidgwick argues that Happiness has to have

Adopting a broader textual view of Sidgwick's notion of pleasure, and its relationship to Happiness, thus reveals greater consistency in Sidgwick's use of hedonism than the chapter in which Sidgwick specifically renounces pleasure as an ultimate end⁴⁶⁰. It is seen that Sidgwick both agrees with Bentham, and departs from him, on the same topic. Either way, even if it is concluded that Sidgwick *is* following Bentham's definition of Happiness, his defence of this notion – and his demoting of the idea of pleasure as an ultimate master – is a far stronger one.

Moreover, we can translate Sidgwick's emphasis on ends and the specific way in which actions that pertain to those ends acquire their own reasonableness directly

some defining features, and this *does* require reference to pleasure; in both *ME* and 'Utilitarianism 1873', Sidgwick is simply describing what utilitarians actually mean when they describe Happiness. What is important to maintain then is that whereas there is a very close connection between happiness and pleasure in Sidgwick's work, as demonstrated by his willingness to continue Bentham's use of the term 'hedonism', pleasure always remains simply a part of what is meant when we say Happiness: It does not appear as a rational ultimate end itself, and is therefore not completely interchangeable with Happiness.

Sidgwick himself further validates this view when he states at I:VII:92 that the term 'Happiness' is indeed ambiguous, and that '...it seems to be commonly used in Bentham's way as convertible with pleasure, - or rather as denoting that of which the constituents are pleasures; - and it is in this sense that I think it most convenient to use it' (I:VIII:92). The narrow distance between Bentham and Sidgwick's positions on pleasure is evident here – but according to the rest of the text, and especially to I:IV, the more accurate view is conveyed by the 'rather' statement, that Happiness is the end, only *comprised* of pleasures. This is particularly evident again in Sidgwick's discussion of Ultimate Good (III:XIV:398-403), where Sidgwick specifically discusses the nature of pleasure, as a separate concept to Happiness, within his broader definition of Ultimate Good as 'Desirable Consciousness'. Bentham, on the other hand, is unambiguous that all terms such as 'good' and 'happiness' can be viewed synonymously, when describing his principle of utility: "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness (all this in the present case comes to the same thing)' (Bentham, *An Introduction*, 1:4). It is this view that Sidgwick rejects.

⁴⁶⁰ In terms of the argument that Sidgwick was a classical utilitarianism, perhaps even more in that line than Mill, Sidgwick's rather inconsistent use of Bentham seems to both support that idea, and challenge it. Sidgwick is certainly more prepared than Mill to use the language of hedonism, and as has been seen here Sidgwick agrees with Bentham's view that Happiness is the end of moral action. In fact, it could be argued that on this point Sidgwick's actually defends Bentham's hedonistic doctrine more efficiently than even Bentham himself did. He may have renounced the ultimate value of pleasure in Bentham's doctrine, but the result is only a far more convincing case for Happiness. But Sidgwick's reliance on Reason, and the fundamental role that he assigns this concept in his view of morality, gives his doctrine far more credibility than those of either Bentham, or Mill.

The personal document (in which it is important to note that Sidgwick does state that he at least began his career in moral philosophy from a utilitarian position) supports the idea that in some ways, there is a considerably wider distance between Sidgwick and Bentham than is historically presumed. It is actually Mill's doctrine that Sidgwick cites as his first utilitarian position, as it gave him 'relief from the apparently external and arbitrary pressure of moral rules which I had been educated to obey' (*ME*: xv). It should be noted that the 'moral rules' in question here were Whelwell's and not Bentham's, but it is still clearly the case that Sidgwick did not consider Bentham's doctrine to be an adequate alternative. This of course immediately then begs the question as to why, if Sidgwick did not consider Bentham's utilitarianism to have been the best alternative to Whelwell and intuitionism, he is specifically working from Bentham in *ME*. The likely answer seems to be that Sidgwick considers Bentham's utilitarianism to be the formal, *academic* version of the doctrine, and given that *ME* is intended to be an academic study of 'the methods of ethics', it is Bentham's doctrine and not Mill's that is the most appropriate to this purpose.

onto the relationship between pleasure and Happiness. Happiness is the rational ultimate end, the pursuit of pleasure/avoidance of pain is reasonable in as far as it pertains to that end. In this way, Sidgwick's hedonism will be an essential part of his normative ethical theory, in which universalistic hedonism is represented by general Happiness.

3.4. Sidgwick's Utilitarianism

Thus far, Sidgwick's utilitarianism has been seen to be classical to some extent (in its partial agreement with Bentham's hedonism), but less so in other ways that are undeniably significant. On the basis of his understanding of rightness especially, and the consequences of that understanding for means and for ends, Sidgwick's approach to the doctrine is particularly distinctive.

It will also be noticed however that at this stage of *ME*, Sidgwick's utilitarianism is still rather vague. There has been no protracted discussion of utilitarianism as a method of ethics, and where apparently utilitarian principles have appeared they have often been bound up with concepts that pertain as much to deontological principles. This is primarily because of the emphasis on the Reason that Sidgwick believes to be common to all methods, and that has clearly appeared as such in his work as explored above. At the beginning of I:VI, the shared appeal to Reason has particular prominence. 'The aim of Ethics', Sidgwick reminds the reader, 'is to systemise and free from error the apparent cognitions that most men have of the rightness of reasonableness of conduct, whether the conduct be considered as right in itself, or as the means to some end commonly conceived as ultimately reasonable' (I:VI:77). He also specifically reiterates that there is a 'threefold difference in the conception of the ultimate *reason* for conduct' – one of which is Perfection, another Duty, and the third Happiness (I:VI:78, *italics mine*)⁴⁶¹. Sidgwick now asks exactly what *are* these ultimately valid reasons for acting. It is in his answers to this question that Sidgwick begins to draw sharper and more familiar distinctions between the deontological and utilitarian positions – between the ends of action as opposed to the execution of duties that are prescribed by unconditional rules – in order that he

⁴⁶¹ See also I:VI:83, where Sidgwick takes for granted that all moral viewpoints include a notion of 'the ultimate reason for doing what is concluded to be right'.

might investigate each of these independently⁴⁶². His statement that the two types of utilitarianism, egoism and universalistic hedonism, ‘agree in prescribing actions as means to an end distinct from, and lying outside the actions; so that they both lay down rules which are not absolute but relative, and only valid if they conduce to the end’ (I:VI:84) provides both the clearest view of his own understanding of utilitarianism, and the most recognisable summary of a standard, characteristically utilitarian position thus far⁴⁶³. From this point, it becomes possible to determine the specifics of Sidgwick’s own understanding of utilitarianism.

This section continues to follow *ME*, mainly chronologically, in order to assemble and assess the relevant material that contributes to completing that understanding of Sidgwick’s utilitarian doctrine. It must be noted from the outset of this part of the discussion that, despite my argument for the deontological properties of Sidgwick’s philosophical intuitions, I will not simply be denying below that Sidgwick is a utilitarian (or, more accurately from the point of view of my overall argument in this thesis, that he is *also* a utilitarian, as well as a *kind* of deontologist)⁴⁶⁴. Sidgwick’s arguments for utilitarianism cannot and should be ignored, and as will be seen, Sidgwick approaches the utilitarian principle in two distinct ways - ‘negatively’, in Book III, wherein he will argue that the principles of Dogmatic Intuitionism are actually found to be subordinate to utilitarianism, and ‘positively’, in Book IV, where he will show how utilitarianism is the systemiser of common sense morality. However, I will still be presenting Sidgwick’s utilitarianism as a very specific type, and, crucially, as in vital relationship with the deontological properties that inform common sense morality. As will become apparent in the following sections (and as is directly concurrent with this thesis’ central claim) it is almost impossible to discuss the role and value of utilitarianism in Sidgwick’s work without reference to the role

⁴⁶² As Sidgwick puts it, “in my treatment of the subject, difference of *Method* is taken as the paramount consideration: and it is on this account that I have treated the view in which Perfection is taken to be...a variety of the Intuitionism which determines right conduct by reference to axioms of duty intuitively known; while I have made as marked a separation as possible between Epicurean and Egoistic Hedonism, and the Universalistic or Benthamite Hedonism to which I propose to restrict the term Utilitarianism” (I:VI:83-84).

⁴⁶³ ‘Utilitarianism 1873’ also presents an extremely condensed account of the doctrine as it appears in *ME*.

⁴⁶⁴ It is my view that a conclusion on Sidgwick’s status as a utilitarian cannot be fully made until the particulars of the mutually informing relationship between utilitarian and deontological principles that he perceives and develops in Book III have been properly expounded. For this reason, the real conclusion on whether or not Sidgwick is a utilitarian naturally appears in chapter 4 of this thesis, in conjunction with the wider claim that he is only a utilitarian on a deontological basis, or vice versa, where the argument of chapter 4 makes it possible to argue this point.

and value that he assigns to deontological principles – and *ME* itself affirms this. Therefore, only a certain stage in Book IV can be reached before the investigation into Sidgwick's utilitarianism requires that we return to his arguments made in Book III – where I argue we see a meeting between Sidgwick's deontological principles, and his utilitarian one⁴⁶⁵. It is that meeting of utilitarian and deontological principles that will then be examined in detail in chapter 4.

3.4.a. 'Difficult Calculations'

This investigation into the nature of Sidgwick's utilitarianism has now reached the third feature outlined in the chapter's introduction, which is that Sidgwick recognises early on the 'defects and difficulties of the hedonistic method'. At I:VI:87, he states that he is intentionally leaving his examination of utilitarianism until last, partly because it seems simpler that egoistic hedonism should precede universalistic hedonism - and partly because it seems as though the claims of the Intuitionist method should be examined before these are compared 'with the results of the more doubtful and difficult calculations of utilitarian consequences'.

It is interesting that Sidgwick should take this view of the utilitarian method, and even more so that he believes that the Intuitive (or deontological) position should be examined first. His saying that it is a more complicated method does not necessarily mean that Sidgwick believes utilitarianism to be inferior to Intuitionism, but there is certainly a sense here that Sidgwick believes utilitarianism can only be adequately understood via *some* kind of reference to, or relation with, Intuitionism. It is also clear from this passage that Sidgwick recognises the difficulties of measuring the outcomes that are integral to the utilitarian method. Both of these factors retain a high profile throughout Sidgwick's investigation into utilitarianism. Chapters II, III and IV of Book II specifically address the difficulties of outcome measurability as it appears in the context of egoistic hedonism⁴⁶⁶. Book IV of *ME* then presents

⁴⁶⁵ Although this non-sequential approach may seem somewhat incoherent, it actually accords directly with the structure of Sidgwick's own approach to utilitarianism, as it develops over Books III and IV. It is in Book III that Sidgwick reaches what he considers to be some inevitable conclusions on the role of utilitarianism; he then temporarily leaves aside these conclusions for IV:I and IV:II, only picking them up again when his own discussion has reached the stage at which it naturally recurses back to those conclusions from Book III.

⁴⁶⁶ Sidgwick rejects the empirical-reflective method of comparing pleasures and pains, and also casts serious doubt on the ability of common sense to do this (II:IV:153-161). These chapters provide a

Sidgwick's exposition proper of the formal model of utilitarianism, understood as referring to general Happiness, rather than private. Here, Sidgwick temporarily lays aside his own conclusions on utilitarianism drawn in Book III in order to define exactly what is meant by 'utilitarianism' - and he does not delay in drawing attention to those difficulties to which he has already alluded. His opening comment is that whereas utilitarianism is supposed to be a position with which we are all familiar, on 'closer examination' it is actually comprised of a number of different theories and concepts that are widely diverse, and in some cases even seemingly unrelated (I:IV:411)⁴⁶⁷. For this reason, Sidgwick gives his clearest – and most obviously classical – definition of the doctrine so far: 'By Utilitarianism is here meant the ethical theory, that the conduct which, under any given circumstances, is objectively right, is that which will produce the greatest amount of happiness on the whole; that is, taking into account all whose happiness is affected by the conduct'⁴⁶⁸. Sidgwick then proceeds to eliminate from the definition some of the most common misperceptions. He distinguishes universal hedonism from egoistic hedonism, arguing that the apparent affinity between these two positions is nothing more than the result of a confusion⁴⁶⁹; he denies the utilitarian theory that holds that the moral sentiments are derived from sensory experience, arguing that this is not adequate proof of the ethical doctrine (IV:I:412-413); he argues that Universal Benevolence should not be understood as providing the only right *motive* for action (IV:I:413). He then turns to the principle

point of reference for later in *ME*, when Sidgwick points to the same difficulties as they appear in general-Happiness utilitarianism.

⁴⁶⁷ Even Sidgwick's own utilitarianism, as it has appeared so far, is an example of how diverse interpretations of this doctrine can be.

⁴⁶⁸ Sidgwick states that the term 'utilitarianism' can also be used interchangeably with 'universalistic hedonism'. This is significant: The latter phrase is not simply another rendering of the word 'utilitarianism' - rather, the two words together embody the key features of his utilitarianism that Happiness is the ultimate rational end, and that this end, being rational, cannot help but convert itself from private Happiness to the general Happiness – which is thus itself established as an end, universal and unconditional. The details of these claims are part of the very core of this thesis' argument, and Sidgwick's argument for universalistic hedonism, which appears in Book III of *ME*, is discussed at length in this thesis' chapter 4. This outline is given here so as to verify that Sidgwick has previously given strong grounds for believing utilitarianism to refer to the Happiness of *all*, and not to the private happiness of individuals, and this outline is required to provide a full understanding of Sidgwick's utilitarianism, with which we are concerned here.

⁴⁶⁹ This makes it acceptable to entirely omit egoism from Sidgwick's utilitarianism (see also Sidgwick's express restriction of the term 'utilitarianism' to 'Universal Happiness' at I:VI:84, II:I:119, and III:I:199). Sidgwick declares that the reality of the difference between egoistic and universal Hedonism is 'so obvious and glaring, that...we seem rather called upon to explain how the two ever came to be confounded, or in any way included under one notion' (IV:I:412). Sidgwick blames Mill for the confusion, and Mill's unsuccessful argument that because every one *does* seek his own happiness, it is an obvious inference that we *ought* to seek the happiness of others. The transition from psychological to ethical hedonism, Sidgwick argues, can actually only be to the egoistic version.

itself, restating what is meant by the term 'greatest Happiness' (that it is to be understood as the greatest possible surplus of pleasure over pain), and asking whose happiness is to be included in the 'all' referred to in that principle (IV:I:414).

The answer to this last question, which Sidgwick gives as 'any sentient being', throws up its own further complications. As difficult as it already is to scientifically compare hedonistic measurements among people, it becomes even more problematic when the interests of non humans are also included (IV:414)⁴⁷⁰. But even if we do only account for the happiness of human beings, there are then the issues of non-temporality, and posterity. Because happiness cannot be affected by the time in which a person lives, it follows that the utilitarian must also be concerned with the interests of future human beings. Ultimately, the prescription of happiness on the whole seems, to Sidgwick, to require that we increase the population to the extent at which the largest number of people possible is sharing the average happiness (IV:I:416)⁴⁷¹.

The method of utilitarianism, also, is beset by difficulties. IV:IV details the problems involved in accepting utilitarianism as a means of determining right conduct, and as a basis from which to gain an adequate system of rules (IV:IV:467). Humanity is so vastly diverse that to decide on general rules would seem impossible. Life is also constantly subject to change, and in the attempts to systemise this by arguing that the 'Preservation of Society' should be the ultimate 'scientific criterion' of

⁴⁷⁰ It is arguments such as these that have caused a revival of interest in the work of Sidgwick among modern day utilitarians, and that point to Sidgwick's enduring relevance to modern ethical problems such as population ethics, foreign state intervention in warzones, and environmentalism. As seen in the literature review, Peter Singer – along with his use of Sidgwick - has specialised in how utilitarian considerations apply to animal welfare. In terms of the distribution of happiness, Sidgwick argues in the absence of a prescribed mode of distribution, the only principle that it is appropriate to adopt is that of 'pure equality – as given in Bentham's formula, "everybody to count for one, and nobody for more than one'. This emphasises the fact that Sidgwick's Happiness is highly egalitarian, which is why Sidgwick and Rawls can often be united on points of political theory.

⁴⁷¹ This is as opposed to increasing the average happiness. To some extent, Sidgwick is arguing for a version of 'the greatest good for the greatest number' here, given that he believes the greatest number of people possible should be existing to share average happiness. But Sidgwick points out something – a problem - that he believes utilitarians to have missed. This is that if utilitarianism demands Happiness *on the whole*, we ought to weigh the amount of Happiness gained by an addition to the population against the amount lost by the already existing population. This results in a situation in which it is not the average Happiness that is maximised, but rather the *number* of people living, who share in the average Happiness (IV:I:416). Derek Parfit has famously referred to this point as 'the repugnant conclusion', as it implies that we ought to prefer a society in which there are very many people living, but average Happiness is very low, to a society in which there are not as many people, but the average Happiness is greater (*Reasons and Persons*, pp.381-387). Parfit's 'repugnant conclusion' is an important aspect of modern day population ethics, and has proved to be something of a paradox, which apparently cannot be solved without causing further unacceptable conclusions. Sidgwick is arguably the fore-runner of Parfit's theory.

moral rules, Sidgwick can find no real justification (IV:IV:470-473). Ultimately, Sidgwick concludes, utilitarianism cannot construct a morality *de novo* – it must also rely on the morality of the existing social order.

This last point will provide chapter 4 with important evidence for my own ultimate argument that utilitarianism must necessarily work with the fundamental deontological properties that informed common sense morality - but here it is evidence of Sidgwick's willingness to observe and admit to the problems involved in utilitarianism as a method of ethics. He is certainly not, in these chapters which specifically concern utilitarianism, simply advocating or defending it. Overall, utilitarianism emerges in Sidgwick's hands as a complicated, strict, and demanding method of ethics, the precepts of which are often paradoxically as incalculable as they are necessary. It is also demanding in its scope, applying as it does to everything from socio-political concerns, to the interests of future generations. Furthermore, as Sidgwick observes, the types of calculation required by the implications of utilitarianism for populations are 'grotesquely incongruous' with our actual ability to execute these calculations in practice, which has an 'inevitable inexactness'.

Sidgwick's response to the difficulties, however, is important for an accurate understanding of what *does* become Sidgwick's own position on the doctrine. This response is not just to acknowledge that utilitarianism is a particularly stringent method of ethics, but to argue that this stringency must actually be observed. He says: 'that our practical Utilitarian reasonings must necessarily be rough, is no reason for not making them as accurate as the case admits; and we shall be more likely to succeed in this if we keep before our mind as distinctly as possible the strict type of the calculation that we should have to make, if all the relevant considerations could be estimated with mathematical precision' (IV:I:416).

This adherence to a rigorous and far-reaching version of utilitarianism might seem to be at odds with my claim that Sidgwick's utilitarianism cannot be understood without reference to the non-consequentialist aspect of Intuitionism. But there is no necessary reason why a stringent view of utilitarianism should not refer to deontological principles. In fact, that a doctrine is strict and comprehensive and yet still refers to principles outside of its traditional philosophical territory only reinforces the argument that this connection is somehow necessary. And this very point presents itself in *ME* - for it is precisely at the stage in *ME* in which Sidgwick's

utilitarianism is the primary focus that Sidgwick's references to the common sense morality of Intuitionism also begin to appear. Most significantly, this connection arises as a result of Sidgwick's approach to the *proof* of utilitarianism.

3.4.b. Utilitarianism as the Supreme Moral Principle? Positive and Negative Proof

It is in the second chapter of Book IV that Sidgwick turns his attention to the long-standing question of the 'proof' of utilitarianism. This proof is, he says, generally required in order for minds to accept the principle of aiming at the general Happiness, rather than at their own (IV:II:418). For Sidgwick however, the question is more profound than this. Whereas common sense morality will not feel it necessary to prove the authority of rules that are already commonly accepted as binding, a utilitarian who disputes that these rules *are* binding is by default claiming that their moral principle is superior to Intuitionism. Utilitarianism cannot simply be accepted as just another moral principle: As Sidgwick argues, "the Utilitarian prescriptions of duty are *prima facie* in conflict...both with rules which the Intuitionist regards as self-evident, and with the dictates of Rational Egoism". Utilitarianism, then, "if accepted at all, must be accepted as overruling Intuitionism and Egoism" (IV:II:419-420) ⁴⁷².

Here, then, we have reached the fourth stage outlined in this chapter's introduction, which is that Sidgwick finally appears here have materialised as a true utilitarian. This is in addition to the fact that many of the 'self-evident' principles of common sense morality were found in Book III to actually point to utilitarianism – and that the Universal Benevolence that was derived from Sidgwick's philosophical intuitions is, in many ways, a utilitarian principle, however deontological the

⁴⁷² This, Sidgwick states, is the real and only method through which an ethical first principle can be proven. It is not enough to exhibit a first principle by reference to other principles, because then those principles become the premises on which the first depends, and it is no longer then a first principle (II:II:419. See also 'Utilitarianism 1873', and 'The Establishment of Ethical First Principles' (in, *Mind*, Vol.4 (1879) 106-111, p.108). Sidgwick believes this to have been the mistake by Mill. In contrast, Sidgwick establishes his own utilitarian principle on what he believes to be an ultimately rational, self-evident principle. Sidgwick's approach indisputably avoids the shortfalls of Mill's, and succeeds in defending utilitarianism more adequately than Mill had done. But the establishment of that utilitarian principle, and its inherent connection to deontological moral properties, is the concern of chapter 4.

foundation of that principle might be⁴⁷³. This is the point at which I do not deny Sidgwick's utilitarian status. But, crucially, this is also the point at which I can begin to point to the *other* conclusion that forms part of the two between which 'the truth lies', in the connection passage from Book III - the part of the conclusion that reads, 'particular duties have still to be determined by some other method'. For a directly comparable passage appears here at the beginning of Sidgwick's discussion of the 'proof' of utilitarianism, that is, on my view, one of the most profoundly important statements of *ME*. It occurs when Sidgwick points out that there is a fundamental paradox that arises in 'proving' utilitarianism in relation to other moral principles. This is that if Intuitionism denies that its own moral principles are valid, thus making the way for the 'proof', then there is actually no need for that proof in the first place. Either Intuitionism believes its rules to be binding, in which case the Intuitionist will continue to believe that they are binding – or they concede that they are not, in which case there is no need for 'proof' of utilitarianism as a first principle. 'Perhaps we may say', Sidgwick concludes in response to this, 'that what is needed is a line of argument which on the one hand allows the validity, to a certain extent, of the maxims already accepted, and on the other hand shows them to be not absolutely valid, but needing to be controlled and completed by some more comprehensive principle' (IV:II:420).

Both of these 'connection' passages are representative of the fact that Sidgwick perceives there to be some kind of essential relationship between common sense morality, and the utilitarian principle⁴⁷⁴. Clearly, the proof of utilitarianism cannot be discussed *without reference* to the principles of Intuitionism, in some form. But we do require to know the nature of that relationship. Sidgwick is, after all, talking about the possibility of utilitarianism being the sole and supreme moral principle, and this needs to be accounted for in my wider argument that Sidgwick's moral theory of

⁴⁷³ See this thesis' chapter 2 for my deontological argument for Universal Benevolence. This derivation of Universal Benevolence as a utilitarian ultimate rational end from a primarily deontological basis is the focus of chapter 4, and will be discussed at length there accordingly.

⁴⁷⁴ I outlined in chapter 2 that I am referring to these two passages as 'the connection passages', on the basis of their similarity, and for the fact that they both, on my view, indicate Sidgwick's overall position that there is a relationship between deontological and utilitarian moral properties. We might call that 'some kind of essential relationship' a 'middle ground' between deontology and utilitarianism, and this is not technically wrong: My argument for Rational Benevolence in chapter 4 can be described as creating a middle ground between the two types of principle. However, 'middle ground' misses the depth, details, and necessity of the relationship, which I will draw out in the next chapter. 'Middle ground' can be used superficially then (as I will occasionally do) but not simply as a substitute for Rational Benevolence.

Rational Benevolence is not, ultimately, completely utilitarian – or that it is at least a very particular kind of utilitarianism. This relationship, outlined in both connection passages, is what Sidgwick himself now addresses. How, he asks, in light of this paradox, is proof of a principle to be ascertained at all? Sidgwick first reiterates his earlier point that if the egoist is to argue that his *own* happiness is good from the point of view of the universe, then it becomes possible (and necessary, on Sidgwick's view) to draw the egoist's attention to the fact that his happiness cannot be any more of an important part of the universal Good than anyone else's (IV:II:421). This was of course part of how Rational Benevolence was established as Sidgwick's normative theory. But Sidgwick observes that this argument is not enough to convince the Intuitionist, as it still does not show Utilitarianism to be the sole or *supreme* moral principle (IV:II:421). Utilitarianism must therefore 'exhibit itself' in two ways to Intuitionism. It must show to the Intuitionist that principles such as Truth and Justice etc. have only a dependent and subordinate value and that those notions are themselves only vague and in need of further determination; and it must show how utilitarianism sustains and supplements the validity of common moral judgements. These two approaches are Sidgwick's 'negative' and 'positive' exhibitions of the proof of utilitarianism. Again, there is a strong impression here in both aspects of this double relation that utilitarianism is simply going to be held up as the champion moral theory. Sidgwick continues at the end of these passages, 'If systematic reflection upon the morality of Common Sense thus exhibits the Utilitarian principle as that to which Common Sense naturally appeals for that further development of its system which this same reflection shows to be necessary, the proof of utilitarianism seems as complete as it can be made' (IV:II:422)⁴⁷⁵.

But to conclude that Sidgwick now simply supersedes common sense morality – and the deontological properties therein – with utilitarianism is, on my view, to miss the very particular nature of this two-pronged relationship about which Sidgwick is talking. For utilitarianism does not simply emerge as superior, and where it does this is only because it was developed out of common sense in the first place. This begins to be seen when Sidgwick immediately refers the reader back to his argument in III:XIII, stating that he has already in that chapter presented the negative relation of

⁴⁷⁵ For an alternative view on Sidgwick's 'proof', William Langenfus ('The Impossibility of Sidgwick's Proof', in, *History of Philosophy Quarterly*, Vol.17, No.1 (2000), 99-120, p.101) presents a convincing argument for why this proof actually fails.

utilitarianism to common-sense morality. There, it will be recalled, Sidgwick argued that the principles of Intuitionism, or common sense morality, are not self-evident, and that in fact they actually often only point to the further utilitarian principle. But it will also be recalled Sidgwick did *not* simply abandon the non-consequentialist, deontological basis of common sense morality. In fact, he was concerned to show what *can* truly be called self-evident in this way⁴⁷⁶. Remaining with the value that he could technically see in those deontological properties, provided they were truly tested for self-evidence, Sidgwick developed his own philosophical intuitions – which I argued *retain* a visibly non-consequentialist and deontological character. These eventually yielded the utilitarian principle of Universal Benevolence – and thus the negative relation is finally resolved and embodied by Rational Benevolence⁴⁷⁷. It is therefore, I argue, through that negative argument that Sidgwick developed the meta-physical connection between Intuitionism and utilitarianism that allows us to understand the positive relation that he now expounds in Book IV⁴⁷⁸. As utilitarianism

⁴⁷⁶ Roger Crisp points out that Sidgwick is referring to the fact that principles of common sense morality have ‘validity to a certain extent’ because they are independent of inference from self-evident principles i.e. they are intuitive only (*Cosmos of Duty*, pp. 212-213). Where I part ways with Crisp on this is that my focus is on the connection between common sense morality and utilitarianism that Sidgwick develops because of the self-evidence that he *does* eventually attribute to Justice, Prudence, and Benevolence. It is my view that this connection, embodied by Rational Benevolence, makes for a much stronger relationship than one between utilitarianism, and raw intuitions.

⁴⁷⁷ J.B. Schneewind, in his in-depth examination of the two proofs, refers to these positions as the dependence (negative) argument and the systemisation (positive) argument (*Sidgwickian Ethics*, pp.264-5). These are more immediately helpful terms than Sidgwick’s own, given that they represent what each part of the argument is actually doing, but Roger Crisp points out that they might actually lead to some confusion (*Cosmos of Duty*, p.213). This is because part of the positive argument is also aimed at showing that it is utilitarianism on which common sense morality *depends*. However, as Crisp also says, Sidgwick himself is not entirely strict about separating the tasks of his negative and positive arguments – and Crisp’s point is indicative of a wider discussion about Sidgwick’s proof, that is important to my arguments made in this chapter and the next. At the centre of this discussion is the confusion over whether Sidgwick’s argument for utilitarianism in Book III is only negative, or whether his argument for Rational Benevolence actually falls into the category of sustaining and supplementing common moral judgments, in which case this is positive relationship territory. I agree that Sidgwick’s argument in Book III does become a positive one, rather than remaining negative – especially when we consider the fact that it is in Book III that Sidgwick *specifically* argues for Rational Benevolence as ‘proof’ (III:XIII:387-388). This means, at least on my argument, that Sidgwick never actually develops the negative relation at all. And this is vital for my overall argument about the interdependent nature of the relationship between utilitarian and deontological principles. See chapter 4.

⁴⁷⁸ As per my point at fn.54 I do not actually believe that this *is* a negative form of proof. Sidgwick may have intended it to be negative, but this is not how it turned out – and this adds to my argument that the connection between utilitarian and deontological principles is *necessary* to the utilitarian principle, and therefore not capable of simply being made dependent, and subordinate.

Sidgwick specifically discusses Rational Benevolence as being the ‘proof’ of the utilitarian principle in Book III, at III:XIII:387. As I have said in the main text, it is my view that despite the connotations of the word ‘proof’, the particular way in which Sidgwick posits this argument is direct evidence of the relationship of the philosophical connection between Intuitionism and utilitarianism, and as this

must necessarily include reference to Intuitionist principles in order to prove itself at all, the foundations of common sense morality are vital - and through them utilitarianism finds its validation. On the other hand, utilitarianism is also that which systemises and regulates common sense morality⁴⁷⁹.

It is not part of my argument that Sidgwick now proceeds to show how utilitarianism fully explains or completes each of the principles (such as veracity and law abiding, for example) denied in Book III. In many of these cases, the various rules of common sense morality simply are indistinct and ambiguous, and utilitarianism does admittedly emerge as superior⁴⁸⁰. But it *is* my argument that the

connection is the very centre of my main argument for the synthesis-specific discussion of that connection, or 'proof', appears in chapter 4.

⁴⁷⁹ There is a degree to which this relationship, as I have described it, may look somewhat like a Cartesian Circle, in which each part of the whole presupposes the other. This point is most relevant to chapter 4, which argues for the synthesis between utilitarian and deontological moral properties and which will have to make sure that synthesis avoids – if it can – becoming a vicious circle of circular reasoning. But the point has become relevant here for the fact that this relationship between common sense morality and utilitarianism is a direct reflection of that synthesis. The problem might be framed like this: Sidgwick relies upon common sense morality to argue for utilitarianism, then claims that utilitarianism is the systemiser of that common sense morality. But how can the proof of something also be referred to in order to establish that which proves it in the first place?

I do think Sidgwick recognised this problem to some extent, albeit in a negative form, when he describes the paradox of proof. He says that as a first principle cannot be fully proven, it must have to emerge as superior to other principles – on the other hand, if the other principles are not taken as themselves valid, then there is no need for the superior proof. But it is my view that the argument Sidgwick then goes on to make about the relationship between common sense morality and utilitarianism does avoid the problem of the Cartesian Circle. This is because Sidgwick's argument is not completely circular – it does actually begin from somewhere, and this is that we must accept the 'validity to a certain extent' of common sense morality *first*. Only then is the utilitarian principle revealed. Utilitarianism, for its own validation, depends absolutely on the other principles remaining valid. So where utilitarianism is thought to systemise and supplement common sense morality, common sense morality must actually retain its own validity in order for utilitarianism to be demonstrated in the first place. In both of these ways of putting it, common sense morality (or the absolute-principle based/deontological aspect I argued for in chapter 2) emerges as a foundation to this connection between utilitarianism and common sense morality (and, as will be seen, to the meta-ethical connection between deontological and utilitarian properties within Rational Benevolence). This, however, would seem to threaten the very core of my argument, which is that there is a [circularly] interdependent and mutually informing relationship between the two types of moral property which can provide a better explanation for and defence of moral ambivalence. This is not the case however. I do argue, as Sidgwick does, first for the deontological basis of morality; without these non-consequentialist intuitions, the rest of morality could not be established, or identified. But what they identify is a second kind of principle, and once this has been revealed, there develops the relationship wherein it can be seen that both properties of morality are necessary. Deontological properties are required in order to justify utilitarian ones, utilitarian properties then validate those deontological properties. Neither is fully understood or justified without reference to the other – and this is what creates the synthesis. See chapter 4.

⁴⁸⁰ For example, Sidgwick famously states that 'the Morality of Common Sense may be truly represented as at least unconsciously Utilitarian' (IV:III:424), and that "Utilitarians are called upon to show a natural transition from the Morality of Common Sense to Utilitarianism...". He also refers to the 'latent Utilitarianism' of Common Sense on more than one occasion, and says that he is seeking 'to represent [Common Sense] as inchoately and imperfectly Utilitarian' (IV:III:IV). Again, Sidgwick's utilitarian status cannot be denied here, and I will address the implications of that for my theory that argues for the equal importance of deontological and utilitarian principles in the thesis conclusion. I

philosophical connection developed in Book III, between utilitarianism and the deontological principles that Sidgwick *does* consider to avoid this vagueness, is directly reflected in Sidgwick's argument for the 'positive' relation that Sidgwick explores in Book IV. Here we can see the nature of the relationship as it appears in practice, with common sense morality providing the basis and utilitarianism operating as the systemiser, and often the confirmation. That is, this is the *practical* demonstration of those passages, which state that 'the truth lies between between these two conclusions', and that 'what is needed is a line of argument which allows the validity of the maxims already accepted....and shows them to be not absolutely valid'. '...where the current formula is not sufficiently precise for the guidance of conduct', Sidgwick says, for example, 'the Utilitarian method solves these difficulties and perplexities in general accordance with the vague instincts of Common Sense' (IV:III:425). He also says that the method of utilitarianism 'supports the generally received view of the relative importance of different duties', and is 'called in as an arbiter' (IV:III:426), and he draws attention to the fact that in terms of how Benevolence appears in Intuitionism and utilitarianism, there is 'no divergence that we need consider' between the systems (IV:III:430). It will be noted that in all of these passages – and in the many other pages in which Sidgwick discusses the relationship between common sense principles and utilitarianism – Sidgwick is not actually denying or erasing the original deontological nature of common sense morality principles⁴⁸¹. His approach is not simply to allow utilitarianism to 'ride roughshod' over deontological values, replacing deontology entirely as a moral theory. Sidgwick's view is merely that utilitarianism is as much a part of the nature of this common sense morality as those original rules. As Sidgwick says himself, 'the Utilitarian argument cannot be fairly judged unless we take fully into account the cumulative force which it derives from the complex character of the coincidence between Utilitarianism and Common Sense' (IV:III:425). Sidgwick therefore unequivocally confirms that deontological rules remain important, as both of the connection passages suggested that they will be. It now remains for chapter 4 to

will outline what that response will be here however, and simply state that even if Sidgwick's system does emerge as a kind of utilitarianism, this does not detract from the nature – or the importance – of the mutually informing synthesis in Rational Benevolence.

⁴⁸¹ See also especially IV:III:441, 444, 453.

expound the exact *philosophical* basis of that connection, as embodied by Rational Benevolence.

3.5. Conclusions on Sidgwick's Utilitarianism: The 'friendly alliance'

There are a further three chapters in *ME's* book on utilitarianism that follow Sidgwick's initial discussion of the relationship between utilitarianism and common sense morality in IV:III. As was said above however, the end of IV:IV and the chapters that follow it are crucial to my argument for the synthesis, and for that reason discussion of those chapters is reserved until chapter 4. Here, a concluding summary of the findings on Sidgwick's utilitarianism is required.

It was stated at the outset of the chapter that Sidgwick's approach to key utilitarian concepts such as rightness, the role of the good, and the issue of 'means to ends' are atypical of a utilitarian – and especially of a theorist considered to be the paradigm of a classical utilitarian, in the line of Bentham and Mill. It was seen that there is an early predominance of the idea of 'rightness' in Sidgwick's work, good being rejected as an adequate guide for right conduct as early as *ME's* introduction. This was followed by Sidgwick's non-traditionally-utilitarian argument that 'right' is what is ultimately rational, or reasonable, and that this creates a notion of rightness wherein it is not dependent on ulterior results. However, although right cannot be an attribute of means, and therefore cannot be completely defined as 'fit to an ulterior end', the 'ought' that arises from right being that which is rational makes 'ought' a categorical imperative, and this applies as much to Happiness as an ultimate end as it does to certain actions. This interpretation of rightness, it was seen, causes Sidgwick to conflict with the principle of utility, and causes his system to emerge as peculiarly both teleological and non-teleological. It was then seen that Sidgwick rejects pleasure as an ultimate end, but retains Bentham's hedonism on the more refined basis that this is defined as the rational, ultimate end of Happiness, of which pleasure is only a constituent part.

In terms of Sidgwick's own utilitarianism, Sidgwick freely admits to the difficulties of the doctrine, as they appear in both its requirements and its method, and his version of the doctrine is clearly a stringent one. Despite this however, utilitarianism must be defined in relation to common sense morality. The positive relation, I argued, is made possible on the basis of the philosophical connection between

deontological and utilitarian principles that Sidgwick has already shown in Book III. The importance of the deontological basis of common sense morality is maintained alongside utilitarianism, and the relationship alluded to by the two connection passages starts to materialise. Finally, Sidgwick's discussion of utilitarianism, and specifically how its *method* might be constructed, brings him to a stage at which it can go no further without crucial reference to both deontological and utilitarian principles. This brings us in an organic way to the point at which the argument in this thesis also can go no further without expounding the relationship between those two moral properties – which will now be the concern of the next chapter.

It would seem as though the ultimate conclusion to be had on Sidgwick's utilitarianism requires a combination of the ideas that on one hand Sidgwick is not a typical utilitarian, and that on the other hand he is. In many ways – such as his more defensible argument for Happiness, his establishment of 'ought' and 'right' in ends as well as certain absolute principles, and his provision of better 'proof' than Mill – Sidgwick is, as so much of the literature holds, a *better* utilitarian than either Bentham or Mill. It could also perhaps be argued that given his early concern to establish 'right' as applicable to end based systems as well as absolute-principle based ones, Sidgwick is actually building an argument for utilitarianism from the very beginning of *ME*. But I dispute that this is the case. It is my belief, and this is supported by Sidgwick's firm denial that he was intending to defend any one ethical position more than another, that Sidgwick approached *ME* as objectively as possible. His conclusions and observations on the identity of the right and the role of Happiness are not the result of an ulterior motive, but rather of careful, systematic, and neutral examination.

In any case, it is already becoming clear how Sidgwick challenges the traditional divide between deontological and utilitarian principles. As was seen in the previous chapter, and emphasised above, Sidgwick does not eradicate the deontological basis of common sense morality, and on the contrary incorporates it as a vital part of his argument for utilitarianism. And this is not the first time that Sidgwick has referred to such a relationship. In I:VI, Sidgwick states that there is actually a greater practical affinity between utilitarianism and Intuitionism than there is between egoistic and universalistic forms of hedonism (I:IV:85). Despite the two forms of hedonism sharing the concept of pleasure, and of actions as means to distinct ends, Sidgwick's rationale for the more unexpected connection is the simple point that moralists who

maintain that Intuitionism is the valid method of moral reasoning agree that general Happiness is the end for which the rules of morality are the best means (I:VI:85). Again, it cannot be denied that utilitarianism appears to be the superior principle here, and Sidgwick reinforces this with his observation that even fierce opponents of the doctrine will eventually be led to 'utilitarian considerations', when attempting to establish the real necessity for moral rules (I:VI:86). But in the paragraph immediately following, Sidgwick describes how when it first appeared in moral philosophy, Utilitarianism was intended to be in 'friendly alliance' with Intuitionism (I:VI:86), the two forms of morality united against Hobbes' 'audacious enunciation of Egoism'. It was not to supersede but to support the morality of common sense, 'against the dangerous innovations of Hobbes' that had led Cumberland to state that egoism was 'the common good of all Rationals', to be the end to which moral rules were the means' (I:VI:86). As in the other places discussed above, Sidgwick is once more challenging the idea that actions are valid only in as far as they conduce to an end. Then there is Sidgwick's clear assertion in this passage that utilitarianism was *not* intended to surpass, or overtake common sense morality. When Sidgwick later on the page considers the idea that utilitarianism was eventually advanced as the superior moral principle, this antagonism too, 'relates rather to theory and method than to practical results'. That is, where Sidgwick recognises the difference in their philosophical approaches, he does not believe this difference to lead to any real conflict in the practice of each type of morality.

Sidgwick may well be a utilitarian then. But I argue that even in his specific discussions of utilitarianism, there is substantial evidence that his utilitarianism is of a very particular type. It is this type of utilitarianism, that is seen to be in necessary meta-ethical relationship with deontological principles, that will now be examined in the last of the thesis' main chapters. It is this type of utilitarianism that will 'afford a principle of synthesis' (IV:III:422).

CHAPTER 4

‘RATIONAL BENEVOLENCE’

SIDGWICK’S ACCOUNT OF THE SYNTHESIS BETWEEN DEONTOLOGICAL AND UTILITARIAN PRINCIPLES

The findings of chapters 2 and 3, which drew out the deontological and utilitarian aspects of Sidgwick’s work respectively, have laid the groundwork from which this, the final chapter of the thesis, will build its definitive argument – namely, that Henry Sidgwick perceived there to be a necessary and inherent relationship between those deontological and utilitarian principles of morality. Sidgwick, I will argue, understood this relationship to be a sort of synthesis, by which it is meant that the components, in this case the two types of moral value, combine to form something that is more meaningful than just a sum of the parts. It falls to this chapter now to complete the argument for this particular combination of deontological and utilitarian principles, which constitutes Sidgwick’s own moral theory, and is represented by Sidgwick’s term, ‘Rational Benevolence’. The chapter will show that Rational Benevolence is comprised of neither deontological nor utilitarian principles alone, but of a reciprocal co-operation between both types of moral property. It is argued that this co-operation, or synthesis, offers a reconciliation between traditionally divided utilitarian and deontological moral principles, which in turn offers a more accurate and helpful understanding of the human moral experience than theories of morality that exclusively support either one position or the other. In this way, I argue that Rational Benevolence offers a deeper understanding of the moral ambivalence that is often experienced in the gravest of ethical dilemmas.

As regards the structure of this chapter, this is the stage at which all the claims made throughout this thesis about Sidgwick’s deontological principles, his utilitarianism, and his perception of the relationship between them, culminate. It is important therefore that the chapter adheres to a structure that maintains a logical and visible progression from what has already been argued in the thesis, and the continuation and expansion of those arguments into the full theory here. For this reason, the chapter begins with a brief analysis of the terms such as ‘interdependent’, ‘mutually informing’, ‘reciprocal’ and ‘necessary’ that have been used throughout the thesis to refer to the ‘synthesis’ that I believe to exist

between deontological and utilitarian principles of morality, and consequently to the relationship that I believe Sidgwick to have expounded in *The Methods*. This is followed by an examination, which relies extensively upon textual analysis, of the degree to which Sidgwick himself believed there to be this kind of synthesis between the two types of principle. The next section then builds the case that whereas Sidgwick states on more than one occasion that it is his intention to remain entirely impartial throughout *ME*, this impartiality can actually be seen to dissolve as Sidgwick finds himself unavoidably drawn to certain conclusions, from which he does actually construct a moral theory.

The main discussion is then taken up by a comprehensive analysis of Book III, in which Sidgwick uncovers the essential link between the fundamental, self-evident and primarily deontological principles of Intuitionism that are Justice, Prudence, and Universal Benevolence, and the equally utilitarian nature of Universal Benevolence. This is the link that is represented by the term 'Rational Benevolence'. At the beginning of this section, the conclusions of the preceding two chapters on Sidgwick's deontology and his utilitarianism are briefly revisited. But this is not simply a repetition of points. The earlier investigations into Sidgwick's deontology and utilitarianism drew attention to the nature and moral-philosophical value Sidgwick attributed to each type of moral principle independently, and also alluded to the areas in which each principle functions to indicate the other. It is specifically this *latter* context in which the respective deontological and utilitarian properties of Benevolence will be discussed here, appearing now in crucial roles within the wider framework of my argument for that meeting between the principles that creates the reconciled and coherent whole that is Rational Benevolence. The argument for this whole occupies the central sections of this chapter.

For further evidence that Sidgwick developed an inherent link between deontological and utilitarian principles, the final section of the chapter turns to the personal document. Here, Sidgwick explicitly describes his reasons for believing there to be no opposition between deontological and utilitarian principles from his own individual, rather than purely academic, point of view. Sidgwick's own conclusion on this relationship is thus used in conjunction with mine drawn at the end of the chapter.

The results of this chapter will be assessed in the thesis' conclusion, where they will appear in context of the Parental Predicament that was used at the start of

the thesis to demonstrate the serious moral ambivalence that often occurs between absolute-principle and outcome based modes of moral reasoning. This will include a discussion of how this thesis envisages the synthesis embodied by Rational Benevolence to apply in practice, and an assessment of its value for explaining and justifying moral ambivalence.

4.1. The Meaning of ‘Synthesis’

In the thesis introduction, I theorised that the most effective way to address the problem of moral ambivalence, which is characterised by the presence of two equally reasonable and yet apparently conflicting moral principles, would be to show that the two types of moral property are actually inherently *connected* via that shared reasonableness. This, I argued there, would reconcile absolute-principle based (deontological) and outcome based (utilitarian) moral properties into a kind of dynamic synthesis, in which the reasonableness of each type both validates and limits the reasonableness of the other, to an appropriate extent. The various terms that I have used to refer to this reconciliation that I claim can be found in Sidgwick’s work have all been expressly selected on the basis of their ability to precisely describe the nature of that relationship. The vocabulary of ‘interdependent’, ‘reciprocal’, and ‘mutually informing’, is intended to convey the concept that each type of principle – the deontological and the utilitarian – is equal in importance to the other, and that they must rely upon each other for their *actuality*, their *recognition*, and the *role* that they play in the human moral experience. What exactly is meant by each of these qualities must now be examined, in order to generate a series of requirements, or criteria, that this chapter must show Sidgwick’s theory to meet, if it is to be called a synthesis.

Beginning with the ‘actuality’ point, what is meant by this is that deontological properties are only *fully* realised and made certain with reference to utilitarian properties, and that utilitarian properties are only *fully* realised with reference to deontological principles. The synthesis presents a co-reliant situation in which neither principle, despite their obvious differences, can be fully conceived of without the other. Whichever principle is taken first, it must automatically indicate the other. This naturally implies that both principles must also have a degree of their own validity, and thus through that relationship the value of each in its own right is also

fully *recognised*. Again, this is regardless of their opposing natures. When these features are combined, and when the relationship is understood positively, the inevitable conclusion is that both types of principle must have a vital place within the system of human morality, and that at least part of the 'place', or *role*, of each can only be understood with reference to the other principle. The converse of that situation – or the negative understanding of the relationship – is that if the two principles are so divergent and yet *both* have value, then the recognition of their respective values will likely cause each position to have to accept certain limits. Taken all together, this series of conditions is what is meant by the expression 'mutually informing', which is ultimately the most effective way to describe overall what is meant by a 'synthesis' that is not just an amalgamation or a fusion of various ideas, but a system in which those ideas exist in a necessary, dynamic relationship.

In terms of what is meant by the relationship being 'necessary', this is on one hand a slight extension of the words just described. If one property cannot be fully realised without reference to another and vice versa – if each property *generates* the other - then it is given that in a system in which there is more than one property, both are indispensable to that system. But the term 'necessary' is also intended to represent something beyond its merely descriptive qualities. The word implies 'requirement', and describing the relationship between deontological and utilitarian principles as 'necessary', or 'required' applies not just to a metaphysical understanding of that relationship – it also lends to the theory of that relationship a certain impact that it will have on moral philosophy in general. This is because stating that a relationship between deontological and utilitarian principles is necessary (in order for either to be fully understood), is the strongest and most direct challenge to the traditional divide that exists between those principles. As I argued in the thesis introduction, this divide is a fundamental problem in moral philosophy, inaccurate as it is in reflecting the reality of the human moral experience – and bringing some reconciliation to the situation via Sidgwick's theory of Rational Benevolence is the thesis' principal aim. How far this reconciliation is successful, both practically and philosophically, will be assessed in the thesis' conclusion – but here it is important that this potential is understood to be conveyed by the use of the

word ‘necessary’ in relation to the connection between deontological and utilitarian moral principles⁴⁸².

4.2. The Possibility of a Synthesis in Sidgwick’s Work

Before it can be argued that a synthesis of the nature just described exists between deontological and utilitarian principles in Sidgwick’s work, it is necessary to assess whether Sidgwick himself allows that such a thing is possible. Ahead of this even, it is also necessary to question whether *ME* actually offers any kind of moral theory at all. It would be unacceptable to try and extract from Sidgwick’s work something that he himself would not consider to be valid. It would also very unlikely be a successful endeavour. The following sections address Sidgwick’s statements that he himself intended to remain impartial in *ME* as regards an ethical position, and take into account the consequences of this for whether a moral theory can be developed from *ME*. But they also employ textual evidence of Sidgwick’s underlying hope – and this is his own word – that reconciliation between the principles may be possible.

4.2.a. Sidgwick’s Intended Ethical Neutrality

We cannot, when studying Sidgwick, ignore his repeated claims about the aim of *ME* that it is an investigation into the methods commonly used by ordinary people in their moral reasonings, and *not* a treatise on how ethics *ought* to be. He is clear about this in the preface to the first edition, stating that he has ‘put aside temporarily the urgent need which we all feel of finding and adopting the true method of determining what we ought to do’, and that ‘all the different methods developed...are expounded and criticised from a neutral position, and as impartially as possible’ (*ME*: vi). He restates

⁴⁸² Using the word ‘necessary’ to refer to the relationship between deontological and utilitarian principles also implies that at the practical level of how morality is lived and experienced, both types of morality are necessary. This should not however lead to a situation in which both are *enforced*. This is precisely not the intention of arguing for the indispensable presence of both types of morality: ‘Necessary’ is not taken here to mean other interpretations of the word, such as ‘compulsory’, or ‘obligatory’. Rather, it is theorised that if both types of principle are needed, or required, for a full understanding of human morality, then *both* types should be more robustly defended. This point will be discussed further in the thesis’ conclusion.

this position in the second preface, in response to the ‘fundamental misunderstanding’ that *ME* is primarily a defence of utilitarianism⁴⁸³.

The fact that Sidgwick is so careful to point out that he should not be understood as advocating utilitarianism over any other method is important for two reasons. The first, which has been pointed out above, is that it supports the argument that Sidgwick’s work is *not* simply a defence of utilitarianism, as is so often assumed. Secondly, it reinforces Sidgwick’s concern to remain himself unbiased towards *any* one ethical position – something that adds further integrity to his investigation and, later, to its results. We see continuing references to this neutrality throughout Book I – first when Sidgwick outlines his plan for *ME* in the introduction proper – ‘The present book contains neither the exposition of a system, nor a natural or critical history...I have attempted to define and unfold not one Method of Ethics, but several...’, and next in I:VI when he is summarising the findings of the previous chapters – ‘What then do we commonly regard as valid ultimate reasons for acting or abstaining? This, as was said, is the starting-point for the discussions of the present treatise: which is not primarily concerned with proving or disproving the validity of any such reasons, but rather with the critical exposition of the different ‘methods’...’ (I:VI:78). But despite the strength of this objectivity, Sidgwick’s impartiality (or, to use Sidgwick’s turn of phrase, this ‘disinterested curiosity’) is clearly meant to underpin the entirety of the investigation. It would therefore seem to follow that *ME* is unlikely to be offering a moral theory: To do this would necessarily mean *ME* adhering to and supporting a certain position. Sidgwick’s own words maintain a close link between his neutrality and the non-normative nature of *ME*: ‘...we may seek to add to the number of these [ethical] systems’, he says, ‘and claim after so many unsuccessful efforts to have at last attained the one true theory’ (I:I:12) – but Sidgwick does not mean to do this. It is not his ‘primary aim to establish [ethical first] principles’, or to ‘supply a set of practical directions for conduct’ (I:I:14). Sidgwick is clearest about this in the final sentences of the introduction: “I have wished to keep the reader’s attention throughout directed to the processes rather than the results of ethical thought: and have therefore never stated as my own any positive practical

⁴⁸³ ‘I find’, Sidgwick says, ‘that more than one critic has overlooked or disregarded the account of the plan of my treatise, given in the original preface...: and has consequently supposed me to be an assailant of two of the methods...and a defender of the third’. Sidgwick also points out that he has drawn attention to the defects of all the moral theories, not just those of the morality of common sense.

conclusions...” (I:I:14). The question becomes then how, in light of this, it could possibly be put forward at all that *ME* offers a moral theory.

4.2.b The Ideas of Reconciliation and Synthesis in I:I

As much as Sidgwick’s declared impartiality casts doubt on *ME*’s potential to offer a moral theory, there are several areas of the text, even at this preliminary stage of *ME*, that support the claim that Sidgwick is to some extent looking towards the idea that these contrasting ethical principles could possibly be harmonised. The most important of these areas is Sidgwick’s own discussion of the very concept of reconciliation, which appears towards the end of the introduction. Here, Sidgwick admits that the idea of reconciliation is attractive. The various methods of ethics have, according to Sidgwick, been presented as ‘alternatives between which...the human mind seems...necessarily forced to choose’ (I:I:12)⁴⁸⁴, whereas it seems to Sidgwick that the human mind naturally strives for consistency: ‘...because men commonly seem to guide themselves by a mixture of different methods, more or less disguised under ambiguities of language...along with these claims is felt the need of harmonising them...’ (I:I:12). There is, Sidgwick is convinced, a clear impulse on the part of human beings towards trying to resolve the problem of apparently valid yet conflicting moral principles. What is more, Sidgwick specifically uses the language of ‘synthesis’. The word appears twice, both times in direct conjunction with the concept of that reconciliation between the varying practical maxims individuals encounter in their moral thought⁴⁸⁵. In just these three pages of *ME*, there is little doubt that

⁴⁸⁴ This comment also clearly demonstrates Sidgwick’s own recognition of the divide between the major schools of moral thought, at which this thesis is aimed on the whole.

⁴⁸⁵ It also appears again later at I:VIII:102, when Sidgwick makes the point that Intuitionism understood as the deontological system of ethics is ‘in need of rational synthesis’, and later in *ME* at III:XIII:374, IV:II:422, 496, and 497. Most of these examples provide sound evidence that to use the language of synthesis in relation to the relationship Sidgwick perceives between utilitarian and deontological principles is legitimate. It should be noted that in some of these passages (those on 496 and 497 especially), Sidgwick is referring to the possibility of a synthesis between Rational Egoism, and Rational Benevolence, the antipathy between which of course causes the Dualism of Practical Reason. But this is not the only context in which Sidgwick uses the language of synthesis, as demonstrated by the other passages. In the introduction in particular, Sidgwick is simply referring to a synthesis of the diverse practical maxims accepted by most people, that he represents with his series of questions that is designed to show that we ‘answer ‘yes’ to all these’, when their answers are actually quite different, and often contradict each other. There are considerable grounds, then, on which to claim that Sidgwick was also imagining a synthesis between absolute-principle based and outcome based principles – and as will be seen, there is also substantial evidence that Sidgwick believed he *had* achieved at least this extent of synthesis.

Sidgwick is using the term 'synthesis' to represent an ideal system, either personal or philosophical, in which those practical maxims are somehow combined harmoniously - in which they are 'brought into clear relation to each other' (I:I:6. See also I:I:12-13).

Sidgwick's confidence in previous attempts on the part of philosophers to create such a synthesis however is low. His opinion is that writers have 'proceeded to synthesis without adequate analysis; the practical demand for the former being more urgently felt than the theoretical need of the latter' (I:I:12). In Sidgwick's opinion, the difficulties of the various ethical treatises have often been overlooked in the interests of explaining morality in this way, with the result that those difficulties will only reappear the more forcefully later on. Overall, Sidgwick says, 'we get on the one hand vague and hazy reconciliation, on the other loose and random exaggeration of discrepancies' (I:I:13). As apparent as the need for reconciliation between our practical moral maxims might be then, Sidgwick's message is that great caution must be applied in attempting it (as his choice to avoid attempting such a thing also implies). In Sidgwick's astute words, 'the clear indication of an unsolved problem is at any rate a step to its solution' (I:I:13).

It is in that very 'unsolved problem', however, that Sidgwick's solution itself might lie. Sidgwick has already identified and named this problem - it is that of the conflict between different principles that are yet apparently *both* reasonable. This situation is for Sidgwick, as outlined above and mirroring exactly that of this thesis, at the very heart of the problems in ethical theory: 'I am myself convinced that this is the main explanation of the phenomenon...we cannot, of course, regard as valid reasonings that lead to conflicting conclusions' (I:I:6. See also I:I:12). It is at this point that it becomes possible, even this early stage in *ME*, to perceive the thread in Sidgwick's thought that will, in spite of his clearly stated intentions, lead to Sidgwick creating his own reconciliation, and his own 'synthesis'. That thread is precisely that of the equally reasonable status of each type of principle – and I argue that Sidgwick's recognition of this does actually lead him to the 'solution' he mentions at I:I:13. Rationality was a recurrent theme for those earlier passages taken from the prefaces⁴⁸⁶. The same theme continues in the introduction (see I:I:3, 5-6, 12-14) -

⁴⁸⁶ Throughout these preface passages, Sidgwick speaks of 'the methods of obtaining reasoned convictions as to what ought to be done' (*ME*: v), of 'how conclusions are to be rationally reached' (*ME*: vi - this phrase appears twice in the exact same wording), and reframes his question as 'what

and both chapters 2 and 3 of this thesis repeatedly drew out Sidgwick's consistent belief in Reason as a common basis to both deontology and utilitarianism⁴⁸⁷. It is clear that for Sidgwick, the moral maxims by which most people abide are united at least by the fact that all are generally felt to be reasonable – by which it is meant, according to Sidgwick's own definition, that they are thought to be rational, unconditional, self-evident, and not pertaining to a still ulterior end⁴⁸⁸. It is this theme that Sidgwick will carry through *ME* and into his own conclusions on the role, value, and mutual relationship of deontological and utilitarian principles

To summarise, one of the points made above needs to be recalled – but it also needs to be posited alongside an observation on the text that crucially works in favour of the argument that Sidgwick does make it possible for a moral theory, and even a synthesis between reasonable and yet conflicting principles, to be drawn from his work. The point to recall is that Sidgwick himself originally intended to refrain from 'expressly attempting any such complete and final solution of the chief ethical difficulties and controversies as would convert this exposition of various methods into the development of a harmonious system' (I:I:13); to ignore this would simply be an incomplete reading of the text, which would lead to the de-contextualisation of Sidgwick's arguments. Not only would this weaken any claims made by this thesis, but it would actually weaken Sidgwick's own arguments themselves: Part of the strength of Sidgwick's argument for the reconciliation between utilitarian and deontological principles is the very fact that he did not specifically intend to pursue or develop such a thing⁴⁸⁹. But the crucial observation which is to be borne in mind

among the precepts of our common conscience do we really see to be ultimately reasonable?" (*ME*: xi).

⁴⁸⁷ In addition to the passages already mentioned in the preceding chapters, Sidgwick's view that Reason is common to all methods of ethics is also evident at I:III:23, I:III:36, I:VI:77, and I:VI:83.

⁴⁸⁸ Sidgwick does acknowledge that not all ends are determined or prescribed by Reason (I:I:8), and that we only require to assume that an end that dictates 'right conduct' is adopted as ultimate and paramount on some basis. But to investigate *this* approach to ethics is, Sidgwick states, untenable. The task is only made manageable when we 'confine ourselves to such ends as the common sense of mankind appears to accept as rational' (I:I:9). But in any case, Sidgwick's own normative definition of rightness being that which is rational supersedes this observation.

In an interesting contrast with the fact that Sidgwick considers this apparent reasonableness of differing principles to be the *problem*, the way in which he includes the language of rationality when he mentions the *conclusions* to which people are drawn in their moral processes does actually suggest that Sidgwick is predicting Reason to be part of the *solution* in that it remains relevant to all the outcomes that will emerge as a result of his investigations in *ME* (see *ME*: v, vi). There is an impression in those early words that they are prophetic of the conclusions that Sidgwick himself will reach by the end of *ME* – that despite differences in method, the most fundamental moral principles have Reason in common.

⁴⁸⁹ I will return to this statement shortly.

alongside this is taken from the sentence immediately following that passage. Here, Sidgwick states that he ‘hopes’ ‘to afford aid towards the construction of such a system’ (I:I:13). This hope is of the utmost importance. As the rest of this chapter will demonstrate, that hope eventually evolved. Sidgwick’s investigation into the methods of ethics led him inexorably to a point at which he was able to make statements based on more than just optimism. And Sidgwick himself was entirely aware of this particular result of *ME*. In an astonishingly overlooked passage from the preface to the *second* edition of *ME*, Sidgwick makes the striking claim that he had found the answer to that timeless question of ‘What among the precepts of our common conscience do we really see to be ultimately reasonable?’ Sidgwick does not expand on that answer there (this is for *ME* itself, and this chapter, to do) – it is what that answer represents to which Sidgwick draws attention: “...it supplied the rational basis that I had long perceived to be wanting to the Utilitarianism of Bentham, regarded as an ethical doctrine: and thus enabled me to transcend the commonly received antithesis between Intuitionists and Utilitarians” (*ME*: xi). The significance of this sentence can hardly be overstated. It contains all at once Sidgwick’s adherence to the importance of rationality, evidence that Sidgwick was not simply a disciple of the utilitarianism of Bentham – and, most importantly of all – a direct statement from Sidgwick himself that he had found there to be a means of superseding and reconciling the traditional conflict between utilitarianism and deontology. It is, essentially, incontestable proof of three of this thesis’ central claims.

It seems therefore that that ‘hope’ for a reconciliation between reasonable but varying moral principles of which Sidgwick speaks in the introduction can legitimately be harnessed – and, crucially, developed; Sidgwick’s own words give permission to do so⁴⁹⁰. We can therefore now follow through on that hope for a solution, and expound how Sidgwick achieved it.

⁴⁹⁰ It might reasonable be asked why, if Sidgwick did recognise this particular aspect of *ME* as the second preface suggests he did, it did not appear more prominently in *ME* itself. The only place in *ME* (until we reach Book III and beyond) that Sidgwick makes reference to the discovery that utilitarianism and deontology share such a vital link is in a passing comment that part of his object in the volume is to ‘expound as clearly and as fully as my limits will allow the different methods of Ethics....to point out their mutual relations’ (I:I:14). Considering that the claim Sidgwick makes in the second preface is a monumental one, it seems strange that Sidgwick would not draw more attention to this ‘mutual relations’ aspect of his take on ethics – or that he did not revise *ME* to incorporate it. The prefaces show that Sidgwick made various other adjustments to subsequent editions, some of them fairly major – so why not this one? It could be theorised that either Sidgwick did not consider the discovery of this relationship to be as important as it actually is, or that he thought a degree of relationship between the methods was obvious from the start. Both of these are made plausible by *ME*- and whichever it is, no

4.3. The Tone Changes: Enquiry vs. Conclusions, and the Dissolution of Sidgwick's Neutrality

It is clear from *ME*'s introduction that Sidgwick did at least consider the idea of reconciliation between moral principles – but these comments on reconciliation and the possible role of Reason do not themselves immediately eliminate the fact of Sidgwick's intended neutrality regarding the various systems of ethics. It is only reasonable to assume that Sidgwick went forwards into the rest of *ME* on the basis of that intended impartiality. In order then for my argument that Sidgwick *did* actually produce a reconciliatory moral theory to succeed, it must be shown that there was some kind of change in the text, that bridges this gap between Sidgwick's original neutrality and his eventual definite argument for a specific relationship between deontological and utilitarian principles. It is my view that such a change does occur, gradually unfolding as *ME* progresses, and that that change is a move on Sidgwick's part away from that emphasis on impartiality.

Before that change is examined, it is helpful to assess quite why it might have taken place. There are two possible reasons for which such a change might have happened. The first is that Sidgwick simply forgot his original intention to remain impartial, and wrote the rest of *ME* insensible of wanting to maintain his 'disinterested curiosity'. This, however, is a very unlikely scenario. Sidgwick was nothing if not fastidious in his approach to *ME* – this is part of what has given the

detriment is caused to Sidgwick's argument for that relationship. He still developed it more thoroughly and with greater insight and aptitude than any other philosopher of that time, and in doing so it seems that Sidgwick did see if not its tremendous importance for practical ethics in general, then at least its necessity for understanding a large part of the confusion between moral principles from which he begins his enquiries in *ME*.

However, it seems to me that the more likely answer to why Sidgwick did not embrace his transcending of the 'common antithesis between utilitarianism and intuitionism' lies in the notorious Dualism of Practical Reason. For Sidgwick, a complete synthesis between all the reasonable principles that an individual is inclined to obey was always going to be impossible for as long as he could not reconcile the fundamental contradiction between self-regarding and extra-regarding impulses. It is a well-known fact that the problem of the DoPR plagued Sidgwick until his final days; it has already been seen that Sidgwick was persuaded to change the final word of *ME* from 'failure', but his self-deprecation on this point is at least partially responsible for the reputation that *ME* still has today, of failing its author in yielding an answer to one of morality's and indeed humankind's most difficult questions. As a result, and ironically also as a result of Sidgwick's frequent lamenting of the point, some of the finer achievements of *ME* are overlooked – both by its readers, and by its writer. The personal document however indicates that later in life at least, Sidgwick was more prepared to embrace the fact that he saw a necessary relationship between deontology and utilitarianism, and that this relationship was important, even if he could not reconcile self-interest to the equation. The importance of this document's role in demonstrating this relationship, and the significance of the point in Sidgwick's life at which he wrote it, is discussed in full below.

volume such enduring fame as one of the best expositions of moral philosophy ever written. It would be entirely out of character – both as an academic and as a person – for Sidgwick to have simply taken a different direction in *ME*, with no particular reason.

The second possibility is that there was an actual reason for Sidgwick moving away in *ME* from that impartiality he was so careful to state in its introduction. This is my position. To expand, it will be recalled that it was said above, in 5.2.ii, that part of the strength of Sidgwick's arguments for that reconciliation lies precisely in the fact that he did *not* purposely set out to pursue it. What is specifically meant by this is that Sidgwick gradually left his original impartiality behind *not* explicitly, intentionally – and not simply in order to argue for utilitarianism - but as an unavoidable result of the conclusions that he saw naturally emerging as he pursued his investigations into the two types of morality. It is the purpose of this section to demonstrate how and where this change away from the originally intended ethical neutrality of *ME* occurs. By tracking the gradual move away from his impartiality that unfolds as *ME* progresses, Sidgwick's own aims in *ME* are brought into alignment with the claim made by this thesis that he did eventually arrive at a moral theory, which in turn sets up the basis upon which the chapter's later sections can argue for that moral theory itself.

4.3.a. I:III – I:VIII – The Early Influence of Reason on Sidgwick's Neutrality

The co-existence of the two key features of Sidgwick's work that were illuminated above - his intention to remain impartial, and the prophetic impression of Sidgwick's crucial faith in Reason - can easily be followed throughout the first two books of *ME*. The text here does still largely accord with Sidgwick's planned neutrality, with Sidgwick examining each moral position in turn, drawing out its place in ordinary thought, and paying attention to both its legitimacy and its difficulties. The tone is, as would be expected, mainly theoretical, and investigative. At the same time, there are various topics on which Sidgwick *does* put forward his own conclusions. This is seen in Book I in his rejection of conformity to God's will as a criterion of rightness (I:VI:80), his similar rejection of 'Nature' as a guide for action (I:VI:80-81), his argument for defining 'good' as something other than 'pleasure' (I:IX:106-109), and in Book II in his statements that whereas the empirical-reflective method does not

appear to be a reliable means of establishing ends in egoistic hedonism (II:III:149-150), it transpires to be the only method available (and even then is still vague and uncertain (I:VI:195)).

Sidgwick is arguably drawing these conclusions for the sake of the clarity that will be needed as his enquiry deepens. They are the results of discussions in which Sidgwick –in his characteristically rational and methodical manner - establishes the outlines and definitions with which he will be working in the investigations proper in Book II, III and IV. They are also processes of elimination through which Sidgwick creates a framework that roughly resembles what the ordinary person might consider to be the subjects and principles most relevant to their moral experiences e.g. egoism is discounted as a real method of *ethics*, whereas the intuitionist view and the utilitarian view are held up as representing the main body of commonly conceived moral doctrines (see I:I:8 ; I:III:36-37 ; I:VI:84-85 ; I:VII:91 ; I:VIII:96 ; esp. II:I:119 ; III:I:200). Certainly many of his conclusions are likely to already be familiar to a reader who has some understanding of moral philosophy⁴⁹¹.

Certain other conclusions however, have a different quality. These are noticeable by the fact that they are stated more assertively, and – crucially – independently of the more familiar conceptions with which Sidgwick is otherwise mainly working. The most notable examples are, as has been seen in this thesis' previous two chapters, Sidgwick's argument for the definition of the right (I:III:26-34), his alignment with Kant on what Reason must entail (I:III:33-35), and his rejection of pleasure as an ultimate end. Where chapters 2 and 3 covered the details of Sidgwick's positions on these topics, the emphasis here is that all of these arguments materialise from what Sidgwick *personally* considers to be the principles and dictates of Reason. This is to say that the conclusions he gives in these areas appear to be statements as to what *ought* to be, as opposed to being concepts or theories about what *is* that Sidgwick is merely exploring as part of *ME*'s general investigations.

Most significantly, as the the first book of *ME* progresses Sidgwick's decisiveness on certain subjects that springs from his reliance on Reason makes it possible to detect a point at which it actually begins to supersede his impartiality.

⁴⁹¹ It is also only reasonable to assume that Sidgwick was going to have to reach a certain amount of definite hypotheses, otherwise the entire enterprise of the book would be pointless.

This starts to become most apparent in I:VIII, with Sidgwick's handling of what is meant by the term 'intuitionism', ahead of his investigation into the Intuitionist method proper that he will pursue in Book III. Where Sidgwick's approach was previously one of enquiry, here there is a noticeable impression that Sidgwick is purposefully setting up *his* view of the Intuitionist position⁴⁹². This is in part due to Sidgwick's perception that 'Intuition' is often used ambiguously, and requires a narrower definition (I:VIII:96-98), which he is here endeavouring to give. Arguably, Sidgwick's definitions of the Intuitionist method as that 'which regards as the practically ultimate end of moral actions their conformity to certain rules or dictates of Duty unconditionally prescribed' (I:VIII:96) and 'that certain kinds of actions are unconditionally prescribed without regard to ulterior consequences' (I:VIII:98) are simply recognisably standard of the non-utilitarian, non-consequentialist-duty method of ethics. But Sidgwick's subsequent categorising of three types of Intuitionism (I:VIII:96-104), leads Sidgwick to state clearly that 'dogmatic' Intuitionism, which seems to him to be the sort generally intended by the term 'Intuitive' morality, cannot in fact provide by itself a complete or adequate system of morality. In a rephrasing of his earlier observation that we cannot regard as valid reasonings which conflict, this inadequacy is due to the fact that the confusion of moral principles and methods that make up the predominantly intuitive human moral experience cannot, if it is confused in this way, be ultimately reasonable. 'It is found difficult', Sidgwick says, 'to accept as scientific first principles the moral generalities that we obtain by reflection on the ordinary thought of mankind...even granting that these rules can be so defined as perfectly to fit together and over the whole field of human conduct...still the resulting code seems an accidental aggregate of precepts, which stands in need of some rational synthesis' (I:VIII:102). For Sidgwick then, what is needed is the philosophical *why* of those moral precepts put forwards by dogmatic Intuitionism. He posits even more specifically that one or more principles must be found to be 'absolutely and undeniably true and evident'. In other words, they must be proven to be ultimately reasonable.

⁴⁹² The impression begins early, with Sidgwick's confident statement that no morality ever existed which did not consider consequences to some extent (I:VIII:96). He then goes on to recognise that hedonism may to some extent be called a form of Intuitionism, but promptly expels it from his own definition (I:VIII:98). Sidgwick's tone is in general more self-assured in this chapter, as compared to the first seven chapters of Book I.

It is in this way that Sidgwick's own philosophic view, which demonstrates his confidence in the role and inexorableness of Reason, rises clearly to the fore in these pages. For this concern with what can ultimately be called undeniably true and evident is actually the outline of the task to which Sidgwick will turn himself, in Book III. Here, then, that change in Sidgwick's intentions for *ME* that was outlined at the start of this section is clearly underway. It is to the completion of that change that the next section now turns.

4.3.b. Sidgwick's Abandonment of Neutrality in Book III

This shift in *ME*'s tone that has been caused by Sidgwick's petition to Reason and his resulting willingness to draw his own conclusions when he had previously stated that this was not his objective, carries directly over to the extended enquiry into Intuitionism in Book III. Here - as was seen in chapter 2 (pp.181-186) - Sidgwick subjects the commonly held principles of common sense morality to a rigorous testing (III:II-X) in which he explores how these principles have come to be, and whether or not they are capable of producing any clear and concise guidance for moral action. The fact that Sidgwick undertakes this style of investigation into common sense morality is admittedly not completely at odds with the aims he stated at the beginning of *ME*: Sidgwick says in the preface to the first edition that *ME* is an exposition into the different methods of obtaining reasoned convictions, and that part of its aim is indeed to be critical (*ME*: v). He was also not wrong when he stated earlier at the end of I:VIII (in a passage that acts as a sort of preparation for his upcoming investigation in Book III) that 'for systematic direction of conduct, we require to know on what judgements we are to rely as ultimately valid' (I:VIII:103) - this is, after all, only practical. Furthermore, it is only fair to Sidgwick to acknowledge that he ends III:I with a reiteration that he is not trying to prove or disprove Intuitionism, only to 'obtain as explicit, exact, and coherent a statement as possible of its fundamental rules' (III:I:216)⁴⁹³.

But this statement does not seem to agree with his points made in the preceding three pages of III:I that we require these moral cognitions to be

⁴⁹³ This is echoed at III:XI:338 when Sidgwick specifically declares again that he wishes it to be observed that he has not introduced his own views.

systemised, freed from error, corrected, supplemented, and assessed in terms of their clearness, precision, and authority⁴⁹⁴. In all of these ways of handling moral cognitions, there must be some actual appraisal involved. It would seem therefore as though there are two forms of epistemology in operation here. The first is a purely detached sort of epistemology which is concerned only with how we come to know moral principles - and this is within the original remit of *ME* as stated by Sidgwick. The second form, which appears in Sidgwick's queries regarding the truth and limits of these principles, is Sidgwick's own, philosophical and Reason-centric epistemological method. In light of this, it can be seen that the key point of departure in Sidgwick's move away from impartiality actually appeared in the preparatory passage on page 103, in Sidgwick's use of the word 'valid'. To 'obtain clear and coherent facts' about a situation is not the same as deciding whether or not they are 'valid' - the former is an observation (which is apparently what Sidgwick first intended), the latter is, obviously, a value judgement – and in these chapters in Book III, Sidgwick is not just exploring how we arrive at the principles on which we place value, he is also passing his own judgement on how far these principles are really acceptable in their commonly conceived form⁴⁹⁵. The appearance of Sidgwick's own epistemological method is substantiated further by III:XI, with Sidgwick's retrospective application of his four criteria for self-evidence to the principles that

⁴⁹⁴ It also does not concur with his statement in the first preface that he is concerned in *ME* 'from first to last, *not* on the practical results to which our methods lead, but on the methods themselves' (vi, *italics mine*). Throughout the entirety of III:I, Sidgwick is assertive in presenting his own views (III:I:210; 211; 212 – 213) – and most of these points on which he is unambiguous are specifically rules on which, according to Reason, we must all agree. The main examples are his statement that 'we cannot judge an action to be right for A and wrong for B, unless we can find in the natures or circumstances of the two some difference...' (III:I:209), his simple yet convincing refutation of Kant on the basis that just because an individual can will that their maxim be made universal, that maxim may still – objectively – be wrong (III:I:209-210), and his arrival at the position that 'we have, in fact, to take up the attempt that Socrates initiated, and endeavour to define satisfactorily the general notions of duty and virtue which we all in common use...' (III:I:215). It is my view that these points are crucial to the argument for Sidgwick's skill as a philosopher, arising as they do from Sidgwick's sophisticated and highly detailed enquiry – but within the context of the current discussion, they are further evidence to support the claim that Sidgwick does, when it comes to Reason, adopt a certain unequivocal epistemological view.

⁴⁹⁵ This is evident throughout III:I-X. See, for example, his discussion on benevolence (III:IV:246 – 262) (especially his conclusion on this chapter: '...we must admit that while we find a number of broad and more or less indefinite rules...laid down by common sense in this department of duty, it is difficult or impossible to extract from them, so far as they are commonly accepted, any clear and precise principles...' (III:IV:262)) – and his discussion on justice (III:V:287, and III:V:293 – 'The prominent element in justice as ordinarily conceived is a kind of equality...but when we have clearly distinguished this element, we see that the definition of the virtue required for practical guidance is left obviously incomplete...'). For further examples see also his similar conclusion on the matter of laws and promises (III:VI:311), and his throwing doubt on the absolute validity of veracity (III:VII:315).

have occupied the previous nine chapters. With most of these principles of course failing to meet that criteria, Sidgwick's own conclusions on this matter now appear frequently. There is little doubt by this stage that the reader is receiving Sidgwick's own individual philosophic view⁴⁹⁶.

It is, then, this second epistemological method that Sidgwick has gradually introduced as *ME* has progressed – and it is this epistemology from which Sidgwick's moral theory develops. The Reason-centric nature of that epistemology is fully revealed here in the conviction with which Sidgwick states the conclusions that it yields. Conclusions that are the result of Reason are by default objective, regardless of the writer's opinion. In this way, it doesn't actually matter if Sidgwick himself denies that he has put forward his own views. By implementing this kind of systematic and scientific investigation into the rational self-evidence of these principles, Sidgwick has committed himself to stating actual decisions on the outcomes, whether he intended it or not⁴⁹⁷.

Based on this protracted and mostly negative investigation into the real self-evidence of commonly held moral beliefs, it is understandable that so many of Sidgwick's critics believed him to be attacking common sense morality, with a view to arguing for the superiority of utilitarianism. It is crucial to the central argument of this thesis however to state that those critics' interpretation of III:II-X is not completely accurate. The purpose of drawing attention to *ME*'s abandonment of impartiality is not to point out that Sidgwick has done this in favour of just one (i.e. utilitarian) position, nor to simply point out that Sidgwick seems to have 'changed his mind' about his aims somewhere along the line. The importance of Sidgwick's shift away from ethical neutrality lies in the very fact that it happens at all. As Sidgwick's position on Reason begins to yield conclusions in I:VIII, and continues to do so with even more conviction in Book III - thus dissolving his original neutrality in the process - the entire tone of *ME* changes. This is what was meant when it was said that Sidgwick seems to be inevitably led towards the arguments that he will now develop

⁴⁹⁶ A few pages later, Sidgwick asks the reader to 'travel with him again' through those findings of the preceding chapters of Book III, this time in order to look at them from the perspective of his criteria for self-evidence, and whether or not they really can be stated as truths (III:XI:343). This also strongly indicates that by this stage in *ME*, Sidgwick has his own point that he wishes to impress on the reader.

⁴⁹⁷ It is of course possible to dispute whether the writer's Reason-based conclusions are actually correct, but this is not the point. The writer presumably believes in the accuracy of the results of their process of reasoning, and it is this that commits them to stating those conclusions.

throughout Book III. In these chapters, the *pace* of Sidgwick's writing picks up, and there is a purposeful quality in his writing that suggests that his ideas - underpinned by the Reason and rationality that he considered to be so fundamental to any valid moral principle – are moving towards something certain, and decisive. This sense of purpose pervades the whole of the chapter entitled 'Review of the Morality of Common Sense', and continues into III:XIII, entitled 'Philosophical Intuitionism', as Sidgwick reaches his pivotal conclusions on the value, place, and role in human morality of both deontological and utilitarian principles⁴⁹⁸. Arguing these points as they do, these crucial chapters of *ME* are essentially a highly condensed version of what has already gone before in the volume, and therefore of this thesis also. It is here that Sidgwick, now at his most epistemologically persuasive, points to both the shortfalls and the value of the deontological position that he has already loosely covered in earlier books in a more exploratory (i.e. less definite) way; it is here that Sidgwick will conclude that deontological principles ultimately yield the utilitarian principle of Universal Benevolence. It is in these chapters, with their new authoritative tone, that Sidgwick will arrive at the conclusion of Rational Benevolence, in which deontological and utilitarian principles are bound together in his normative moral theory that answers Sidgwick's 'hope' for such a system. Sidgwick abandoned neutrality and 'disinterested curiosity' not in order to support one type of morality over another, but because of an inherent connection between them which he saw unavoidably emerging before him. It is to that inherent connection that this chapter now turns.

4.4. The Synthesis: Sidgwick's Moral Theory of Rational Benevolence

This chapter has now arrived at a point at which it is possible to argue for Sidgwick's theory of Rational Benevolence as a means of reconciling deontological and utilitarian principles in such a way that they demonstrate themselves to exist in a

⁴⁹⁸ A particularly noticeable example of Sidgwick being now purposeful, rather than speculative comes at the beginning of III:XI when we see that Sidgwick actually considers the task of investigating whether it is possible to find 'real ethical axioms' of real clearness and certainty' through philosophical intuitionism (III:XIII:373) to be something of a *duty*. 'We conceive it as the aim of a philosopher, as such, to do somewhat more than define and formulate the common moral opinions of mankind...his function is to tell men what they ought to think, rather than what they do think; he is expected to transcend common sense in his premises, and is allowed a certain divergence from Common Sense in his conclusions'. We could easily excuse this on the basis that he is right that a philosopher can transcend commonly held ideas – but this was not the intention of *this* philosopher!

mutually informing relationship that can be thought of as a synthesis. It is the purpose of these sections – which form the last and most crucial stage of the thesis – to draw out where this relationship emerges, and precisely how the two principles cooperate within the normative framework of Rational Benevolence. The aim is to bring together and build on those features of Sidgwick’s work – his valuing of the deontological position and his particularly nuanced utilitarianism – that have already been established by this thesis, and that Sidgwick now makes explicit in Book III.

The first section of this stage revisits the results of the investigations into Sidgwick’s deontology and his utilitarianism respectively, in order to re-emphasise precisely these aspects of his work that will now be seen in his theory of Rational Benevolence. This is followed with an extensive analysis of Rational Benevolence as a moral theory, as it appears in III:XIII. The deontological and utilitarian properties of Rational Benevolence are now specifically examined from the perspective of how the two types of principle appear *together*, and, specifically, how they each function to contribute to the theory of Rational Benevolence by remaining deontological/utilitarian but necessarily referring to each other to form a coherent whole. Finally, this specific way in which the deontological and utilitarian components are combined will be subjected to the criteria of synthesis given above. Ultimately, it is argued that this particular connection between deontological and utilitarian moral properties resolves the apparent conflict between them, and thus also removes the long-standing divide.

4.4.a. Reiteration of Chapters 2 and 3

In chapter 2 it was argued that the Intuitionism which Sidgwick associates with the absolute-principle or duty based method of ethics represents a recognisably non-consequentialist form of deontology – and that Sidgwick himself follows a central tenet of that position in his own interpretation of what is meant by ‘right’. This interpretation is that ‘right’, or ‘ought’, pertains to actions that are ultimately reasonable, without reference to ulterior results. On this basis they are unconditional – or non-consequentialist categorical imperatives. Sidgwick concluded that the principles of common sense morality do not admit of self-evidence, and therefore cannot by themselves provide a complete system of ethics, but that his own principles of Justice, Prudence, and Benevolence do emerge as self-evident, given

that they are ultimately rational, and thus right without reference to anything further. As such, the Universal Benevolence that is built out of the rational relationship between these axioms incurs an unconditional 'ought', and also becomes a non-consequentialist imperative. It was thus concluded that this traditionally deontological aspect of morality is, for Sidgwick, integral to the human moral experience, and to a system of normative ethics.

The way in which Sidgwick's Universal Benevolence acquires its rightness leads conveniently on to the findings of chapter 3. For it was noted there that, in accordance with my claim that Sidgwick's utilitarianism is not always typical of his time, his reliance on Reason for his notions of 'right' and 'ought' also applies to the establishment and adoption of certain ultimate, rational ends – specifically that of Happiness. This Happiness is understood by utilitarianism to be the general Happiness, which indicates the principle of Universal Benevolence. It was also seen here that Sidgwick's interpretation of utilitarianism as sustaining the current general validity of common sense morality, (i.e. the positive proof) depends on the philosophic relationship Sidgwick created as a result of *not* dismissing the deontological basis of common sense morality, but rather pressing that basis for full validity. This 'negative proof' (although it is not my position that this *is* actually negative proof as Sidgwick intended it) allows for the continued independent value of deontology/common sense morality, alongside that of utilitarianism, as opposed to simply pointing to the superiority of utilitarianism. It was then observed that as this is the case, we require to know that philosophic basis of that relationship, in which deontological and utilitarian moral properties are *both* sustained and validated at the meta-ethical level. This is the stage of the thesis' ultimate investigation to which we now turn, as these deontological and utilitarian aspects of Universal Benevolence are examined for their combined roles within Sidgwick's moral theory of Rational Benevolence.

4.4.b. The Maxim of Rational Benevolence: Sidgwick's Moral Theory

The construction of Sidgwick's moral theory of Rational Benevolence has already been partly seen, in chapter 2 where I argued for the deontological *nature* of the philosophic intuitions that inform it, and in chapter 3, where it was confirmed that Universal Benevolence is indeed a utilitarian principle. In order to now fully examine

the philosophic connection between these deontological and utilitarian properties however, the construction of Rational Benevolence must be briefly restated. This is partly because the thesis' ultimate focus has now come to rest exclusively on Rational Benevolence as a whole, and partly because we need to clarify that whole so that we can assess the respective *functions* of the deontological and utilitarian properties that have been identified within it.

It is in section 3 of III:XIII (with Sidgwick having made that transition away from asking what we *do* do to asking what we *should* do, that I have been demonstrating above) that Sidgwick finally turns his full attention to developing his own moral theory. He first gives that pivotal statement that whereas 'it would be disheartening to have to regard as altogether illusory the strong instinct of Common Sense that points to the existence of [self-evident] principles', the truth lies between the two conclusions that there *are* absolute principles, but they are so abstract and universal in scope that we require another method for determining particular duties (III:XIII:379). As I have outlined previously, the most significant thing about this passage is its indication that this theory is going to involve, in some way, a combination of principles. Sidgwick then immediately proceeds to the preliminary development of that theory.

First, Sidgwick reintroduces the Kantian-influenced principle of Justice, negatively rephrased in such a way that it avoids Kant's maxim's collapse into subjectivity and is thus fully self-evident. After Sidgwick's alteration, this principle (which reads that 'it cannot be right for A to treat B in a manner in which it would be wrong for B to treat A, merely on the ground that they are two different individuals, and without there being any difference between the natures or circumstances of the two') still evokes the Kantian emphasis on universalizability (III:XIII:380). However, whereas the self-evident truth of this statement itself is clear, the principle is not, on its own, complete as a system of guidance. It provides the framework of how ethical actions are to be managed, but no information on what qualities the actions within it are to have. Sidgwick therefore continues on to point out that this principle was established on the basis of the similarity of individuals that make up a Whole. This Whole is represented by the notion of the 'Good on the Whole' of an individual (III:XIII:381), which is comprised of a temporally neutral consideration of the integrant parts or time frames that are realised throughout conscious life – a consideration represented by the principle of Prudence. This formula of reasoning through which Sidgwick has

arrived at the notion of individual ‘Good on the Whole’ is crucial for how he establishes the next and most critical stage in his moral theory, which is his movement from the notion of individual good to Universal Good. As an individual’s good is comprised of the comparison and integration of all different goods, so Universal Good is formed ‘by comparison and integration of the different goods’ of all individual human – or sentient – existences (III:XIII:382); and as it is only *rational* for a person to aim at their own good on the whole, the Universal Good, by the same logic, must also be an end at which it is rational to aim⁴⁹⁹. If we ‘ought’ to aim at our own whole good because this is only rational, then we *ought* to aim at Universal Good, given that this too is rational. This, Sidgwick states, is also a self-evident principle, and he summarises it thus:

‘...the good of any one individual is of no more importance, from the point of view of the Universe...than the good of any other; unless, that is, there are special grounds...and it is evident to me that as a rational being I am bound to aim at good generally...not merely at a particular part of it’ (III:XIII:382).

From these rational intuitions (that the good of no one individual is worth more than the good of any other, and that we as rational beings are bound to aim at good in general), Sidgwick obtains the maxim of Universal Benevolence, which is that ‘each one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed’ (III:XIII:382). The combination of the ultimately rational and self-evident basis of this principle, with the Universal Good that has been logically derived from individual good, produces the term ‘Rational Benevolence’. This is, essentially, Sidgwick’s moral theory.

Considering the importance of Rational Benevolence, it arrives into the pages of *ME* with surprising brevity. The entirety of the theory’s construction takes up little

⁴⁹⁹ It is conducive here to recall and emphasise from chapters 2 and 3 the point that Sidgwick specifically argues for the notions of ‘ought’ and ‘right’ in I:III:25-35 in such a way that the Reason inherent in these terms produces a categorical imperative of moral obligation: Sidgwick is using that same basis here. Bearing in mind his own negatively stated version of the categorical imperative that ‘it cannot be right for A to treat B in a manner in which it would be wrong to treat B, merely on the ground that they are two different individuals...’ we see forming on p.382 a coherent whole of certain various parts of Sidgwick’s arguments so far, in which the idea of rational ends is combined with the rational notion of ‘ought’.

more than half a page – just two paragraphs⁵⁰⁰. It should also be noted that Sidgwick himself does not expressly state that this is his moral theory. In fact, he offers little explanation of exactly what he means by Rational Benevolence at all; he simply capitalises both words, and places the term alongside ‘Justice’ and ‘Prudence’, which are the only other principles that have a self-evident element. This does suggest however that the conclusion of Rational Benevolence is so obvious to Sidgwick that no further explanation is needed.

But there is no doubt that Rational Benevolence *does* constitute a moral theory. It gives a framework that establishes a universalising test for all moral actions, and then establishes what *kind* of actions they are to be⁵⁰¹. Sidgwick also uses the terms ‘morally bound’, and ‘ought’ to denote the type of relationship that is to exist between these principles and any acting agent (something that is also implied in the very derivation of Rational Benevolence in the first place). This is, undoubtedly, a theory as to how morality *ought* to go – as Sidgwick predicted it would be at the beginning of III:XIII.

4.4.c. The Deontology of Rational Benevolence: Deriving Utilitarianism

Having established the formulation of Rational Benevolence, we can now progress to examine the roles of its respective deontological and utilitarian properties. The function of deontological principles for Rational Benevolence will be examined first; this is followed by an examination of the function of utilitarian principles, and these two investigations together lead directly to how the two types of property are interdependent.

If we follow the progression of Sidgwick’s ideas as they have been presented from the beginning of Book III, it might seem as though Sidgwick is, via his philosophical intuitions, simply offering a solution, or an alternative, to the problem of non-self-evidence that he has been exposing in dogmatic Intuitionism. It is true that

⁵⁰⁰ This is partially because Sidgwick has laid the groundwork for its components (i.e. the links between self-evidence, Reason, rationality, and ‘ought’) earlier in *ME*, and is now able to work more freely with the concepts here.

⁵⁰¹ It is also possible to actually imagine what Rational Benevolence would look like normatively, in practice. Within the conditions that it can never be right for one individual to treat another in a way that it would be wrong for the other individual to treat the first, we are then to treat all people in such a way as we aim at their good on the whole, and that we can universalise that action. I will return to the applicable normative value of Rational Benevolence in the thesis conclusion.

Sidgwick presents his three principles as self-evident (i.e. ultimately rational) where dogmatic Intuitionism failed to do this, but to see these philosophical intuitions as merely a means of Sidgwick supplanting and superseding deontological principles with his own principles (which, when it *is* thought that this is what Sidgwick is doing, are presumed to be utilitarian) is, in my view, to produce an error that then leads to two further errors. The initial error is that this state of affairs would simply demote deontology below utilitarianism. The first of the two resultant errors is that this is then to miss the fact that these philosophical intuitions are actually largely deontological in nature. The second resultant error, which is the most problematic, is the lack of recognition that these deontological properties must and do provide the rational foundations of Sidgwick's entire moral theory, as I will now illustrate⁵⁰².

It was seen in chapter 2 that as Sidgwick's principle of Justice is self-evident and therefore ultimately rational, it is *right* in Sidgwick's sense of rightness, which is *not* 'fit to an ulterior end'. That is, I argued that Justice is a non-consequentialist principle. It is not simply right with reference to ulterior results – it is right absolutely, categorically, and unconditionally, as per Sidgwick's earlier conclusions on these notions. It is also a universalising principle, laying down equality between individuals as a moral dictate. The next axiom, that of Prudence, is also a non-consequentialist principle – and this principle, crucially, establishes *impartial* good as a self-evident and rational end in Sidgwick's theory. This impartial good is obviously supported by the equality requisite from Justice. Prudence and Justice's co-operative argument for 'Good on the Whole' then rationally leads to *Universal* Good, and thus Benevolence. This is to say that Benevolence is *only established on the basis of a combination of these two prior principles*, which I have argued are deontological in nature. It is precisely in Sidgwick's move from individual Good to Universal Good that we see the crucial contributions from both the principle of Justice and the principle of Prudence to the philosophic structure that supports Benevolence. Justice provides a framework of universalising equality; Prudence inserts Good as a rational end into that framework. In the end result of this particular combination, which is that as it is irrational to aim at one good over another in individual life then it must be irrational to

⁵⁰² By 'problematic' here, I am referring to the idea that not recognising this deontological basis obviously poses an obstacle to recognising the possibility that Sidgwick developed a synthesis between deontological and utilitarian moral properties. It is my view that Sidgwick's deontology *has* been largely missed, and that this has indeed prevented his theory from being seen as anything other than utilitarianism, albeit a sophisticated form.

prefer the good of one individual over the good of other individuals, good is thus converted into an impartial and universal form. It is only because Justice and Prudence *are* established as ultimately rational and self-evident that Universal Benevolence – and thus *Rational Benevolence* – is established at all.

What is most crucial about the deontological properties of Justice and Prudence then is that they *remain* deontological, even as Sidgwick progresses to establish the ultimate tenet of his theory, which is commonly thought to be utilitarian⁵⁰³. What this ultimately means is that Sidgwick *derives* his utilitarian principle from a sequential linking of Justice and Prudence, which have deontological foundations. Or – more concisely – it means that Sidgwick derives the utilitarian principle from a deontological viewpoint. If those principles of Justice and Prudence were to lose their deontological status i.e. if they were somehow to be seen as merely consequentialist, or not right in themselves and simply as *means* to a further end, they would lose the crucial characteristics that together logically reveal Benevolence, and Sidgwick would not be able to establish the rational grounds for Universal Benevolence in the first place. Moreover, because the Universal Good (of Universal Benevolence) *has* been established as part of this rational process, it too can finally be established as a rational, ultimate end, and as such it conforms to Sidgwick's definition of rightness because it is not simply made right via reference to ulterior results. The end of Benevolence is therefore *also* a non-consequentialist end; having been established on a self-evident and ultimately rational basis, it retains that

⁵⁰³ Sidgwick does state at III:XIII:386 that 'the self-evident principles laid down do not specially belong to Intuitionism in the restricted sense which, for clear distinction of methods, I gave to this term at the outset of our investigation. The axiom of Prudence [is] implied in Rational Egoism...Again, the axiom of Justice or Equity as above stated 'that similar cases ought to be treated similarly' – belongs in all its applications to Utilitarianism as much as to any system commonly called Intuitional: while the axiom of Rational Benevolence is, in my view, required as a rational basis for the Utilitarian system'. Sidgwick's last comment here is, I believe, of the utmost importance for understanding the relationship between the deontological and utilitarian properties in Rational Benevolence, and for that reason it is discussed independently below. As regards the other principles, it is not detrimental to my arguments – either for those principles as having a deontological basis, or for the deontological/utilitarian relationship – that Sidgwick believes these principles to also apply to utilitarianism. In the fact that Sidgwick deems it necessary to make this point in the first place, he is indirectly recognising that they have been developed out of the Intuitional method, with Intuitional method (i.e. deontological) characteristics. Also, just because these principles can apply as much to utilitarianism as they do to Intuitionism, this does not necessarily threaten their deontological basis, which I have argued comes about as a result of Sidgwick's view of non-consequentialist rightness. Finally, that they *are* applicable to both systems of ethics, and that Sidgwick recognises this, only further supports my argument that Sidgwick's self-evident principles have vital characteristics that allow them to support both deontological and utilitarian moral properties.

primary deontological characteristic⁵⁰⁴. This rationally derived Universal Good is of course more conveniently described with the term 'Rational Benevolence'.

In light of this argument for the deontological basis of Rational Benevolence, deontology may appear to be the dominant moral theory at this stage. We are certainly still missing a crucial part of a 'mutually informing' relationship between utilitarian and deontological principles. It should also be noted that at no point does Sidgwick say that utilitarian principles are the answer to the failings of common sense morality, or that utilitarianism is simply the better alternative theory – this is not the philosophic approach to the solution that he takes, as I have tried to explain above, and in the previous chapters⁵⁰⁵. Having said this, Sidgwick does indisputably enter utilitarian territory when he converts the rationality of pursuing individual Good into the rationality of pursuing Universal Good: The 'point of view of the universe' is easily interpreted as representing a version - albeit rather more refined and better supported one – of the utilitarian premise of the greatest good for the greatest number. The axiom of Benevolence itself is also inherently to do with outcomes – a point that is reinforced by Sidgwick's language of 'good' and 'happiness', which are

⁵⁰⁴ To call Benevolence a 'duty' then, is to say that it is a duty in the specifically deontological sense of that word, in that it is prescribed unconditionally, and without reference to consequences. This is supported by Sidgwick's use of the language 'morally bound'. Obedience to Benevolence is not optional, it is a categorical imperative. Of course, it has already been seen that Sidgwick also deems ultimately rational *ends* to be categorical imperatives (which is a more traditionally utilitarian line), but in both cases, as I have argued, both ends and certain duties are *only* established on the basis of Sidgwick's definition of 'rightness' being that which is ultimately reasonable, and therefore unconditional, categorically prescribed, and right without reference to ulterior results.

On the topic of the 'duty of Benevolence', this phrase appears in Sidgwick's admittance that he had previously 'observed that the duty of Benevolence as recognised by common sense seems to fall somewhat short [of self-evidence]' (III:XIII:382), but he resolves this issue here by arguing that this new maxim of Benevolence is to be understood as operating within the context of Universal Good, which gives a practical *ought*, and gives this general rule that the good of no individual can be held to be more important than the good of any other. This enables Sidgwick, by way of further evidence that supports Benevolence as a self-evident principle, to draw on the life observation that 'the plain man' will 'unhesitatingly' agree that it would be morally wrong for him to pursue his own happiness if it involved sacrificing the greater happiness of someone else (III:XIII:382). Sidgwick may be guilty of employing some optimism here however; somewhat uncharacteristically, he gives no grounds for this statement about the plain man's agreement with this form of Benevolence.

It is also interesting that Sidgwick is so confident that people will agree that the happiness of other people cannot be sacrificed for the sake of their own, given his perpetual struggle with the apparent reasonableness of Rational Egoism. It could be argued that to be an egoist does not necessarily entail actually sacrificing the happiness of others, but Sidgwick has still always had trouble with the fact that it is very difficult to convince an Egoist that he ought to pursue the happiness of others at all. Still, it remains my view that in his argument for Rational Benevolence, Sidgwick resolved the DoPR in favour of this position (as against Rational Egoism) more effectively than he himself believed.

⁵⁰⁵ The specific language of utilitarianism is noticeably absent from the passages in which Sidgwick establishes the maxim of Benevolence.

of course predominantly utilitarian themes⁵⁰⁶. I do not disagree with the general statement that utilitarianism is present in Rational Benevolence – this point is of course vital to my argument that deontological principles are so inherently bound up with utilitarian ones that one principle cannot be taken without the other, and the function of utilitarianism is discussed presently. To this same end however, I can now add the final and most substantial support to my long-running claim in this thesis that Sidgwick's moral theory of Rational Benevolence is *not wholly utilitarian*. This support appears in the fact that, as just explained, utilitarianism is derived from a deontological basis. Although utilitarianism itself does acquire a self-evident status, as will be seen, Sidgwick does not construct it *ex nihilo*. It depends entirely on those prior deontological principles for its own endorsement – and Sidgwick himself confirms this. In a direct reflection of my point that the utilitarian principle is derived from the framework that creates Rational Benevolence, Sidgwick points out that 'the axiom of Rational Benevolence is', in his view, 'required as a *rational basis* for the Utilitarian system' (III:XIII:387 – italics mine). This indicates that very function and role of the deontological properties of the self-evident principles that I have been trying to make clear above – and also indicates that Rational Benevolence is not simply 'utilitarianism'. Sidgwick is clearly drawing a degree of separation between them in that passage, and this is because utilitarianism depends on Rational Benevolence being established in the way that it is. Sidgwick's move to utilitarianism may be a rational and logical one, but it is not a simple or explicit transition.

Utilitarianism is not, then, immediately granted the level of supremacy that might be expected of the purely utilitarian theory that Sidgwick's is thought to be⁵⁰⁷. Rather, it emerges here, almost casually, at the end of Sidgwick's rational-intuitional

⁵⁰⁶ The fact that the dictates of Benevolence are to do with *aiming* at this particular end does not affect my argument for Benevolence *itself* being a non-consequentialist end. As I have said previously, by 'non-consequentialist' I do not mean that the actions Benevolence requires are not consequentialist; they clearly are, given that they are aiming at that end of Benevolence. What is meant by this is that the nature of the *maxim* of Benevolence is non-consequentialist, because as a rational, ultimate end it is not itself right with regard to ulterior results. As a moral principle, Benevolence is absolute, and unconditional.

However, it was also seen in chapter 3 that Sidgwick's attributing this kind of rightness to certain ends does give to the means to those ends a certain degree of 'ought', given that it is only reasonable to do what pertains to an end recognised as ultimately reasonable. This is, I argue, one of the principal achievements of Rational Benevolence, and one of its finest triumphs – the way in which it incorporates both fixed deontological principles and outcome based utilitarian ends, under the same definition of rightness, and reasonableness.

⁵⁰⁷ It is for this reason that I suggest 'Rational Benevolence' is more accurate way of describing Sidgwick's moral theory, as opposed to 'utilitarianism'.

epistemological argument. It is at least certainly clearer here than anywhere else in *ME* that Sidgwick's intuitionist epistemology, crucially informed by an appeal to Reason, extends to both deontological and utilitarian moral principles (see III:XIII:386-387), and this is of course important to my wider argument for the link between them⁵⁰⁸. But I conclude this section with the emphasis that the function within Rational Benevolence of the rational deontological properties inherent in Justice and Prudence is to logically establish and then support and endorse the maxim of Benevolence. For the purposes of building a mutually informing synthesis, the next stage of the argument assesses where utilitarian properties also are indispensable to the establishment of Rational Benevolence.

4.4.d. The Utilitarianism of Rational Benevolence: Contextualising Deontology

I outlined above that Sidgwick is more plainly utilitarian once he has established Universal Good, and the maxim of Benevolence, as the ultimately rational end of moral action. The end of Universal Good is, of course, a distinctly utilitarian notion – and it is this that Rational Benevolence ultimately validates, and represents⁵⁰⁹. Sidgwick, too, is clear that utilitarianism essentially defines his moral position. Having just stated that Rational Benevolence is required as the rational basis for the utilitarian system, Sidgwick follows this lead and declares, in the very next line, that 'Accordingly, I find that I arrive, in my search for really clear and certain ethical intuitions, at the fundamental principle of Utilitarianism' (III:XIII:387). On the next

⁵⁰⁸ At III:XIII:388, Sidgwick refers to 'the intuition of Rational Benevolence'. Now that Sidgwick has claimed intuitionism with a small 'i' as part of his own moral epistemology in these pages, the distinction between this type and Intuitionism as representative of the deontological method of ethics has largely disappeared. This further supports my argument in chapter 2 that there is a close relationship between Sidgwick's Intuitionism with a capital 'I', and Sidgwick's own epistemological intuitionism – and that this is represented by the philosophical construction of Rational Benevolence, which includes deontological-Intuitional elements, and results in utilitarian intuitions. Ultimately, this emerging dissolution of the boundary between Intuitionism and Sidgwick's own epistemological intuitionism demonstrates the dissolution in Sidgwick's work of the boundary between deontology and utilitarianism.

⁵⁰⁹ When 'Good' is interpreted to be Happiness, as Sidgwick interprets it, and when that Happiness includes the notion of pleasure, the result is actually a hedonistic form of utilitarianism. (This is where, as seen in chapter 3, Sidgwick qualifies as a Benthamite utilitarian). But Sidgwick maintains the focus on Happiness. At the end of III:XIII, Sidgwick exchanges the language of 'good' for 'happiness' (III:XIII:388), and proceeds in III:XIV to explain his reason for this (III:XIV:391-407). Sidgwick concludes in that chapter that desirable consciousness can be the only ultimate good, and that this can quite simply be described as Happiness (pp. 401-406). This fully explains why Sidgwick refers with ease to 'Happiness' rather than 'Good' as a rational, ultimate end throughout *ME*, and I have therefore followed Sidgwick in doing this in these sections.

page, he makes a similar statement: 'Utilitarianism is thus presented as the final form into which Intuitionism tends to pass, when the demand for really self-evident first principles is rigorously pressed' (III:XIII:388).

This utilitarian aspect is strengthened by Sidgwick's implementing Rational Benevolence as the 'proof' of utilitarianism, which is usually taken in the literature to quite literally be proof that Sidgwick's theory is wholly utilitarian. This proof is something that he does not consider Mill - who had specifically attempted to give to Bentham's doctrine this authoritative aspect in which it was thought to be in need - to have provided (III:XIII:387-388). Sidgwick's issue with Mill's argument is that it establishes neither an 'ought', or even a rational end at which we reasonably ought to aim⁵¹⁰. His own theory of Rational Benevolence, on the other hand, logically establishes the Universal Good, or Happiness, as the rational end, and as an immediate result of this also establishes the maxim of Benevolence as a directive rule of reasonable, *right* conduct that as such incurs an 'ought'. It is this rationally derived rightness that Sidgwick considers Mill to have failed to provide, and that Sidgwick believes he has constructed via arguing for Rational Benevolence as a rational and therefore categorical duty. In Sidgwick's words, '...there is a gap in the expressed argument, which can, I think, only be filled by some such proposition as that which I have above tried to exhibit as the intuition of Rational Benevolence' (III:XIII:388).

These conclusions on the utilitarianism of Rational Benevolence might seem to throw considerable doubt on my argument that there is a mutually informing relationship between deontological and utilitarian principles in that moral theory - and on my claim that Rational Benevolence is not entirely utilitarian. They certainly do not seem congruent with the conclusion I have just drawn in the preceding section about

⁵¹⁰ Sidgwick outlines how Mill reaches the idea that general Happiness is a good to the aggregate of persons via the prior idea that each person's happiness is a good to that person. Sidgwick also points out that Mill argues that the only evidence that something is desirable is that it is desired, and that this is the only proof that is required as far as the general Happiness is concerned. Sidgwick parts ways with Mill most significantly on the next point, related to the previous, which is that Mill has argued that the utilitarian principle is a 'standard of right and wrong', and a 'directive rule of conduct' (III:XIII:388). Therefore, by stating that the general happiness is what is desirable on that particular utilitarian understanding, Mill *must* mean that we are morally bound to aim at it - that is, that we *ought* to aim at it. But a simple aggregate of actual desires does not necessarily constitute an actual desire for the general happiness - there is no logical or substantive way in which this desire can, to use the language of what is sought here, be 'proven'. If there is no actual desire on the part of an individual (or a group of individuals) for the general happiness, then there is not only a lack of an ought - there is simply nothing at which we reasonably *ought* to aim in the first place.

the importance of deontological principles. I maintain however that both of those claims can still legitimately be made.

Firstly, the fact that the utilitarian principle *is* eventually established as self-evident is actually in itself crucial to the full formulation of Rational Benevolence. It was seen in chapter 3 that ultimate, rational ends are also 'right' as per Sidgwick's definition, as they are not simply fit to some further end. It was then argued above that Universal Good *emerges* as an ultimately rational end of that kind, on the basis of the rational, self-evident principles that inform it. More specifically, once Universal Good has been made self-evident through that process, it emerges as *the* ultimately rational end – and once this has been established, Benevolence then assumes its own role. This is that it retrospectively *binds together* and collates the other two foundational principles. This, then, is how the function of the self-evident utilitarian properties of Rational Benevolence becomes apparent. To expand on the need for this function, Sidgwick has already mentioned that the principle of Justice is not by itself adequate for the construction of a full system of ethics - and Prudence, by way of its own argument for individual good, must logically point to a further good, i.e. the *general* good: The addition of Benevolence as the definitive moral maxim to the formula of Sidgwick's moral theory resolves these shortfalls, by giving those principles full *context*, as part of a comprehensive moral theory. It is most specifically this contextualising capacity that is the function of utilitarianism within Rational Benevolence. This, I emphasise, does not mean that utilitarianism is somehow revising those two principles, or that they have now become mere *means* to the further end of utilitarianism; each principle is still fully independent in terms of its own self-evidence, and they are both still 'right' in a way that does not involve any results that they might bring about. As has already been seen, Justice and Prudence retain their deontological status because without that basis the utilitarianism created by their logical, sequential linking collapses. Once Universal Benevolence has been derived however, those deontological principles do retrospectively become crucial components of a utilitarian theory.

This specific contextualising and validation of Justice and Prudence by the utilitarian principle is thus the first way in which a co-operative relationship between deontological and utilitarian properties is demonstrated within Rational Benevolence. But I said there are two ways in which I believe it is possible to defend the claim that Rational Benevolence depends on both self-evident deontological principles *and*

self-evident utilitarian principles – and this second approach is more comprehensive. It involves my arguing that an alternative reading of that text in which Sidgwick discusses Rational Benevolence as the *proof* of utilitarianism is possible. This is a reading that, unlike the standard one outlined above, does not allow Rational Benevolence to be simply subsumed within the utilitarian doctrine. This particular reading of Rational Benevolence as utilitarianism's 'proof' is what fully reveals the equal and *simultaneous* necessity of both deontological and utilitarian principles, and thus brings together all of the arguments made above into the final argument for the mutual relationship between them.

4.4.e. Rational Benevolence as the 'Proof' of Utilitarianism: The Equal Necessity of Deontological and Utilitarian Principles

Two things have emerged from these respective investigations into the deontological and utilitarian properties of Rational Benevolence. One is that the theory's original foundations are deontological. The other is that its ultimate identity is utilitarian. My reading of Sidgwick's proof incorporates and links these two factors by demonstrating that a) deontological foundations are key, b) that they retain their deontological status, c) that the utilitarian principle can be established as an ultimate rational end without this eliminating the deontological properties, and d) that they are engaged in a mutual and interdependent co-operation. This is demonstrated by assessing two ways in which proof presents itself in Sidgwick's work – the first as a way of establishing the true *rightness* of a moral theory, and the second as Sidgwick's proposed 'double relation' between utilitarianism and the deontological basis of common sense morality.

4.4.e.i. Rational Benevolence as Proof of Utilitarianism's Rightness

The very fact that Sidgwick was looking for *proof* of utilitarianism in the first place should in itself give pause to the conventional line of thought that Sidgwick was purely a utilitarian. For what, exactly, it may be asked, is meant by this term 'proof'? What does it mean to prove something? Proof is definable as 'confirmation', 'verification', or 'substantiation' – to prove something means to find evidence that establishes a fact, or the truth of something. To be the 'proof' of something,

therefore, is no small role. In Sidgwick's words on the lack of proof for utilitarianism in Mill's doctrine, he describes the 'need of some such procedure to complete the argument' as being 'very plain and palpable' (III:XIII:387) - and it is not insignificant that Sidgwick establishes the rational basis of Rational Benevolence *before* he reaches utilitarianism. Sidgwick makes it clear that he has been searching precisely for a vital *rational* basis for the utilitarian principle, that is, of something to qualify it as *right*, on logical and reasonable grounds⁵¹¹. To utilise the correlation created by Sidgwick between the terms 'rational' and 'ought', utilitarianism requires the *rational* so it may become *right* in the sense of without reference to ulterior results, and incur a true categorical *ought* – the absence of which was the reason for his criticism of Mill's argument⁵¹².

This 'proof', then, allows Sidgwick to establish the utilitarian principle as an ultimately rational end - and such a rational end is, according to Sidgwick's definition of rightness, right unconditionally, and not on the basis of its consequences. It is thus clear that this 'proof' must retain a deontological character, even when it is brought into such close connection with utilitarianism – as was said above, if Benevolence did not have the sort of rightness in which it is established as a maxim that is right categorically and without reference to ulterior results, then Rational Benevolence would no longer provide the appropriate rational basis or 'proof' for utilitarianism, for which Sidgwick was searching⁵¹³. This understanding also accords with those synonyms for 'proof' just given, such as 'verification', or 'substantiation'. It further confirms the idea that utilitarianism is only truly validated by deontological principles⁵¹⁴.

⁵¹¹ It has already been noted that Sidgwick leaves the investigation into utilitarianism until the final Book of *ME* (I:VI:87). There, Sidgwick does not state that he is leaving utilitarianism until the end because he considers that doctrine to be in want of a real basis, but he *does* use that language in III:XIII, at 387-388.

⁵¹² Even if the proof passage is still read as conveying the notion that Rational Benevolence merely fills a gap in the wider, superior notion of utilitarianism, that gap is the rational 'ought' – and 'ought' is fundamental to any morally binding system of normative ethics, as Sidgwick effectively suggests in his rejection of Mill's attempt at proof. On this reading also then, Rational Benevolence remains independent of utilitarianism, and indispensable to it.

⁵¹³ Sidgwick's choice of words is also important. Proof, he says, is 'required' (III:XIII:387) – that is, it is necessary for the confirmation of the utilitarian principle.

⁵¹⁴ This point runs parallel to the argument made at the end of chapter 3 that the method of common sense morality provides a foundation from which the method of utilitarianism can proceed. Both arguments pertain directly to the 'middle ground' between the two methods that Sidgwick alludes to in the connection passages.

On the other hand, if something *is* proven, and thus validated in this way, then what is proven also receives a degree of authority itself - and this is also extremely important. Again as I have shown above, once utilitarianism is recognised via the combined rationalities of Justice and Prudence that establishes the ultimate rationality of Benevolence, it assumes its own self-evidence and thus its own role in the full realisation of Rational Benevolence. Without the Benevolence that Justice and Prudence both help to generate, these prior principles are themselves left unsubstantiated. Overall, this combination – represented by this reading of the word ‘proof’ - of deontological properties as the basis of Rational Benevolence and utilitarian properties as its result, is a direct demonstration of the true significance of Sidgwick’s appeal to Reason to establish the rational and moral rightness of *both* deontological and utilitarian properties. Neither deontological nor utilitarian properties lose their fundamental rightness within Rational Benevolence, and neither need do so. Even more significantly, neither can be allowed to lose their particular rightness. Rational Benevolence as the rational basis – the *proof* - of utilitarianism is confirmation of Sidgwick’s ability to simply transcend that traditional deontological/utilitarian divide, in which only one or the other of these types of moral property is right⁵¹⁵. Both deontological and utilitarian properties are equally necessary, in their respective ways, to the full realisation of Rational Benevolence.

4.4.e.ii Negative and/or Positive Proof

Despite my position on the particular and valuable nature of utilitarianism’s ‘proof’ just laid out, it would be wrong to ignore the fact that Sidgwick himself also discusses the issue of proof at length. This was seen in the previous chapter, where Sidgwick’s twofold relation of utilitarianism to Intuitionism brought the chapter to a close

⁵¹⁵ It can now also be seen that Sidgwick’s theory of moral obligation, as represented by Rational Benevolence as a normative moral theory, includes *both* the notion that certain actions are right without reference to ulterior results, and the idea that certain rational *ends* are also right in this non-consequentialist way – as was predicted would be the case in chapter 2. The demands of Justice and Prudence are right without reference to consequences – but so too is the maxim of Benevolence, or Rational Benevolence. On the other hand, the dictates of Benevolence are, of course, consequentialist in nature. They only incur the ‘ought’, however, by extension of Sidgwick’s argument for what is right being that which is ultimately rational, even when this is an end. Where non-consequentialist rightness is the predominant form in Rational Benevolence, ‘means’ are, to some extent, also incorporated.

because I argued that the *positive* version of utilitarianism's proof that Sidgwick develops in Book IV of ME cannot be fully understood without first understanding what he deemed to be the *negative* version that he gives in Book III. Sidgwick himself referred the reader back to Book III for that negative proof, and I followed suit, arguing that the negative proof provides the philosophic connection required for that understanding, and thus completes the relationship to which Sidgwick is referring in the two connection passages. This chapter has already been developing that philosophic connection, and here specific attention is drawn to how it informs both of Sidgwick's versions of proof. The first thing of note however is that the form of proof as it appears in Book III for which I have just argued is actually slightly different to the form of proof that Sidgwick himself considers Book III to have provided. 'Negative' proof stipulates that other principles are found to have 'only a dependent and subordinate validity' to the principle of utilitarianism - and it will be immediately clear from my preceding arguments that this is not what I believe Sidgwick to have 'proved' in Rational Benevolence⁵¹⁶. This is, I argue, because Sidgwick actually does something different in Book III to that which he claims he has done. Sidgwick does *not*, in Book III, only 'negatively prove' utilitarianism, as per his own definition: The utilitarian principle of Benevolence may emerge as the rationally ultimate principle, but it still remains that Sidgwick *derived* this principle from *other already self-evident* principles. And where Universal Benevolence may indeed bring together and contextualise those other principles, as I have argued it does, they are not made subordinate to it, or dependent upon it (at least, not in the way Sidgwick means; I argue for a degree of dependency from Justice and Prudence towards Benevolence, but only as part of the interdependent relationship in which Benevolence is also dependent on Justice and Prudence)⁵¹⁷. Sidgwick did not simply abandon the non-consequentialist, deontological basis of common sense morality. Rather, he maintained and *developed* it. It is the result of this that offers the philosophic connection that I argued in chapter 3 was present in Book III – and *this*

⁵¹⁶ Nor does the sort of negative proof as Sidgwick describes it yield the kind of 'philosophic connection' between deontological and utilitarian principles that I stated in chapter 3 would provide the overarching 'truth' that 'lies between two conclusions' theory that represents the two connection passages.

⁵¹⁷ It is true, however, that the general rules of common sense morality, such as Truth etc. *were* shown to be lacking in self-evidence, and thus dependent on some further principle, which Sidgwick of course establishes to be Rational Benevolence. It is Sidgwick's own philosophical intuitions that I argue are not merely made subordinate, or dependent.

type of connection would seem to accord more accurately with Sidgwick's definition of the 'positive' relation - in which 'utilitarianism sustains the general validity of the current moral judgements' (IV:II:422)⁵¹⁸. This connection then gives us grounds on which to explore the full range of ways in which the positive proof appears in Book IV – and this exploration is illuminative of the fact that both common sense morality and utilitarianism are, in particular ways, present, and active.

It should be noted that Sidgwick's conclusions in IV:III ('Relation of Utilitarianism to the Morality of Common Sense') are initially drawn mainly in favour of utilitarianism as the superior principle. He states, for example, that the morality of common sense may be truly represented as at least unconsciously utilitarian (IV:III:424, see also IV:IV:463), that utilitarians are called upon to show a natural transition from common sense morality to utilitarianism (IV:III:425), that utilitarianism solves the vagueness of common sense morality (IV:III:425), that common sense is 'inchoately and imperfectly utilitarian' (IV:III:427, see also IV:III:453, 454), and that there is a latent utilitarianism in common sense (IV:III:438). These are all clearly examples of Sidgwick remaining true to what he projected the 'positive proof' of utilitarianism would be – that it shows utilitarianism to systemise, cohere, explain, and define the imprecise rules of common sense morality. But this is *not* Sidgwick's ultimate conclusion on the entire relationship. It has already been seen in this thesis' previous chapter that Sidgwick acknowledges the difficulties of utilitarianism as a *method* of ethics, and in IV:IV, where Sidgwick discusses this, he concludes that it is not possible to 'frame with adequate precision a system of rules, constituting the true moral code...as deduced from Utilitarian principles' (IV:IV:467)⁵¹⁹. His central

⁵¹⁸ This also corresponds with the confusion over whether Sidgwick really is establishing 'negative proof' in Book III. I suggest that my argument for the incorporation of both deontological properties (left over from the investigation into dogmatic Intuitionism, or common sense morality) and utilitarian properties into Rational Benevolence might suggest a reason as to why this confusion occurs, even if it cannot resolve it (as I agree that Sidgwick is not consistent on what he apparently considers to be negative proof).

⁵¹⁹ He also adds that utilitarianism as a system is 'beset with serious difficulties' (IV:IV:467), that is seems 'prima facie absurd to lay down a set of ideal utilitarian rules for mankind generally' (IV:IV:468), that humanity is too diverse for any one set of utilitarian rules to apply universally (IV:IV:468,469), and that it is not possible to construct 'a perfect form of society', which might determine how all action ought to go (IV:IV:470).

Sidgwick clearly reveals himself to be a cultural relativist in these passages. He could perhaps even, on these words, be called a moral relativist. But this would seem to be too extreme. If nothing else, my argument for the deontological basis of Sidgwick's moral theory (of Rational Benevolence) has shown that Sidgwick does adhere to an understanding of rightness which yields unconditional duties. However, his awareness that moral norms may not be completely universal in a practical sense actually supports my argument that Sidgwick's Rational Benevolence offers a more accurate, comprehensive and ultimately workable moral theory than those that argue for moral absolutism. This

summary is that ‘the utilitarian, in the existing state of our knowledge, cannot possibly construct a morality *de novo* either for man as he is...or for man as he ought to be and will be. He must start, speaking broadly, with the existing social order, and the existing morality as a part of that order...’ (IV:IV:473-474). The end product of the ‘positive proof’ for utilitarianism being held alongside the enduring importance of common sense morality in this way is Sidgwick’s statement at the beginning of IV:V that we must accept common sense morality as generally utilitarian, but also accept that the morality of common sense is the ‘actually established machinery for attaining this end, which we cannot replace at once by any other’ (IV:V:475). In examining what this system might look like Sidgwick presents a fine balance between utilitarianism and common sense morality, in which a utilitarian would clearly conform to the ‘the Positive morality of his age and country’, and endeavours ‘to promote its development in others’ (IV:V:475). The term Positive Morality represents Sidgwick’s view that ‘though the imperfection that we find in all the actual conditions of human existence...is ultimately found even in Morality itself...still, practically, we are much less concerned with correct and improving than we are with realising and enforcing it’. He continues, in clear terms, ‘the Utilitarian must repudiate altogether that temper of rebellion against the established morality...into which the reflective mind is always apt to fall when it it first conceived that the established rules are not intrinsically reasonable’⁵²⁰. To paraphrase

is because where the deontological properties of Rational Benevolence maintain a certain degree of absolute rightness, the utilitarian properties allow for a degree of flexibility, that, on Sidgwick’s understanding of utilitarianism, are still ‘right’ provided that they remain rational, or reasonable. There is a crucial balance between the two types of principle, in which they both justify and limit each other. This is precisely the sort of theory that I argue is necessary for situations such as the Parental Predicament, and I examine the potential of Rational Benevolence to provide it in the thesis’ conclusion.

⁵²⁰ When considering the relation of utilitarianism to the moral judgements of common sense, Sidgwick says, it is convenient to begin with the idea of duty – that to which a man is bound and obliged – as this is the ‘more important and indispensable of that morality (IV:V:480). Sidgwick is examining here how the utilitarian might respond practically if he concludes that a different rule would be more conducive than the rule supported by common consent. But it is significant that Sidgwick recognises that aspect of common sense morality to be the more important. For he argues that the utilitarian will also feel the effect of breaching those common rules (IV:V:482). If someone is to accept a new moral principle, Sidgwick argues, it is likely that his attachment to the previous ‘habits and sentiments’ would decay and disappear, and that they would only seem to have relevance in as far as they are supported by the sympathy of others. But, Sidgwick points out, this sympathy of others is actually hugely important, as this is where individual moral impulses acquire a ‘large part of their effective force’. It is difficult to measure exactly how and where such considerations will apply to a utilitarian, but the point is, this constitutes ‘an important rational check upon such Utilitarian innovations on Common-sense morality as are of the negative or destructive kind’ (IV:V:483). This idea of a ‘rational check’ on utilitarianism’s possible modifications of common sense morality is a further example of ways in which Sidgwick maintains its vital place, alongside the very particular

Sidgwick, utilitarianism may never truly be able to supersede the role of common sense morality – and, further, that it perhaps *ought not* to do so. Despite Sidgwick's discomfort with the idea of enduring moral conflicts, when it comes to the practical level of lived morality it must be acknowledged that the two methods of ethics share some vital ground.

It will be noticed that my own argument for the relationship between Intuitionism, or common sense morality, does not focus on the practicalities of how the two methods work together, but rather states that Sidgwick has used the same deontological *basis* as Intuitionism for his own moral theory, by identifying which of its principles can be taken as self-evident, and therefore ultimately reasonable, and right. I still maintain this; it is not common sense morality per se that meets with utilitarianism. Sidgwick's argument for the practical relationship is still important however, for the fact that it demonstrates a situation in which utilitarianism validates and sustains the deontological basis of the latter, but in such a way that this dynamic

utilitarianism that he developed in Book III. Sidgwick also mentions that utilitarianism will more likely be required to enforce old rules, rather than introduce new ones, and that when it does come to the conflicting codes it may actually be best if there are two opinions as to a course of action, as it may well be right that A should do x, while B,C, and D blame him for it (IV:V:491). This last point poses an interesting challenge to my argument in this thesis that moral ambivalence would be more effectively managed if *both* courses of action in an ethical dilemma are seen to be simultaneously *right* and reasonable, rather than allowing divergent views to continually conflict. Sidgwick, here, appears to stand in direct disagreement with me, first by maintaining that two different kinds of conduct cannot be right under the same circumstances, and then by arguing that divergent opinion may actually be expedient. At this point I simply have to disagree with Sidgwick, and this is a result of my arguing that the equal reasonableness of both deontological and utilitarian principles within Rational Benevolence results in a normative theory which specifically legitimises more than one course of action (how I envisage this to look in practice is addressed in the thesis conclusion). Sidgwick's own view of the Kantian maxim however, which he states again on p.486 that he accepts as self-evident, emerges in comparison as almost *more* traditional than even my deontological argument for it renders it to be, in that Sidgwick adheres to its dictate more rigidly than my theory does. The question therefore becomes how is it possible to argue for that non-consequentialist and self-evident principle as providing the foundation of Sidgwick's moral theory, which makes that theory both deontological and utilitarian, and then waive it in some way when even Sidgwick himself does not. My answer to this is that I see Sidgwick's theory of Rational Benevolence differently to perhaps how Sidgwick did. It is my view that Rational Benevolence *does* have a recognisably deontological basis, and that this combined with the utilitarian principle creates the synthesis for which I am arguing overall, and that this could produce a moral theory that normatively allows for both absolute-principle based and outcome based moral approaches to be justified. Sidgwick, it perhaps must be admitted, did not himself develop *this* particular understanding of Rational Benevolence, at least explicitly. But I do believe that it is a legitimate reading, and I believe that the rest of *ME* – i.e. Book IV and the argument therein that I am drawing out in the main text - demonstrates the very particular nature of Sidgwick's utilitarianism, that is represented by Rational Benevolence. Rational Benevolence is, therefore, *my* reading of Sidgwick's theory, and my and Sidgwick's divergence over the issue of whether two courses of action can both be justified shows this up. But this does not mean that Rational Benevolence *itself*, in the way that I have described it, is not available from Sidgwick's argument, and that it cannot be used in the exact way that I propose. It is still the case that Sidgwick's construction of the theory offers the particular reading that I have developed.

is sometimes limited, and the direction of dependency actually reversed. In this way, I argue, ‘positive proof’ is a direct reflection of the argument that I have presented above – that the utilitarianism in Rational Benevolence, however prevalent, *depends on prior principles that have been developed out of the basis of common sense morality* (even if that method’s principles themselves are inadequate). That was the nature of utilitarianism’s proof for which I argued above, and Sidgwick confirms it here when he admits that principles outside of utilitarianism are also indispensable. Such a positive relation, he says, ‘affords a principle of synthesis, and a method for binding the unconnected and occasionally conflicting principles of common moral reasoning into a complete and harmonious system’(IV:II:422). That synthesis simply is not complete without common sense morality – utilitarianism by itself cannot provide a complete account of how to construct a system of ethics. And it is this point that I argue can be found in the construction, and in the ‘proof’, of Rational Benevolence.

Finally, Sidgwick closes *ME* with a chapter that he entitles, ‘The Mutual Relations of the Three Methods’. Having completed his ‘pretty full examination of the mutual relations of the Intuitionist and Utilitarian methods’ (496), Sidgwick concludes that:

‘We have found that the common antithesis between Intuitionists and Utilitarians must be entirely discarded: since such abstract moral principles as we can admit to be really self-evident are not only not incompatible with a Utilitarian system, but even seem required to furnish a rational basis for such a system’ (IV:496)

Here, Sidgwick has returned to the concept (including the exact same wording) that the only Intuitionist principles which we can admit to be self evident provide this *rational basis* for the utilitarian system – and he reaffirms that utilitarianism can only be fully comprehended on the basis of those Intuitionist principles that can truly be called self-evident – as per my argument above. This particular reading of Sidgwick’s use of Rational Benevolence as the rational basis, or proof, for Utilitarianism can thus challenge the more familiar claim that this is merely evidence of Sidgwick’s endeavour to validate the utilitarianism that he believed to be to the supreme moral principle⁵²¹. I agree, as I have stated, that Rational Benevolence does indeed

⁵²¹ This reading, and also these passages, also demonstrate Sidgwick providing significant evidence that he believes there to be no divide between deontological and utilitarian principles.

validate utilitarianism, but the alternative reading of Sidgwick's proof holds that it is not simply the role of Rational Benevolence to validate utilitarianism, and then disappear into the wake of Sidgwick's eventual grand argument for utilitarianism as the superior moral theory. Utilitarianism itself relies on initial identification and endorsement by the deontological properties of Rational Benevolence. Both types of moral property are essential – because both share the crucial common ground of being ultimately reasonable, and through this they are logically united in such a way that each has a role in the validation and sustaining of the other.

If any further 'proof' is needed to support this second reading, Sidgwick provides it himself in the preface to the second edition of *ME*. Here, even having been directly challenged on the point, Sidgwick *still* does not concede to his critics that he is defending Utilitarianism over common sense morality (*ME*: x). All that Sidgwick is definite about is the conclusion that the argument above has itself drawn out, that the rational basis that informs the absolute-principle based, common sense morality position also informs the utilitarian position, and thus unites them: 'I ask myself', he says, 'What among the precepts of our common conscience do we really see to be ultimately reasonable?...The answer that I found to it supplied the rational basis that I had long perceived to be wanting to the Utilitarianism of Bentham, regarded as an ethical doctrine: and thus enabled me to transcend the commonly received antithesis between Intuitionists and Utilitarians (*ME*: xi).

4.4.f. The 'Truth' Between These Two Conclusions': Deontological and Utilitarian Principles as Mutually Informing in Rational Benevolence

In my argument for Rational Benevolence as the proof of utilitarianism, I have tried to show that both types of moral property are necessary in order for that proof to be fully realised. I have also argued that this proof is not simply the one-way validation of utilitarianism, but actually a more interdependent affair in which this 'proof' must also retain its deontological properties in order to function in that role. This relationship, I believe, is a development of the relationship to which Sidgwick alludes in the connection passages, when he states that Intuitionistic self-evident maxims must be allowed on one hand, while another method is required to systemise them. It will hopefully have become clear by this stage that, according to this argument, the different moral properties in Rational Benevolence are engaged in an interdependent

relationship, and that both have a role in supporting and substantiating the other. This might well be already called a mutually informing relationship, but as this relationship is the main focus of this chapter, it will still be conducive to purposely examine the system, and its specifics.

The main tenets of the relationship can be stated quite easily. Justice and Prudence have a deontological basis in that they are self-evident and ultimately rational, and therefore right without reference to anything further. Through a logical sequential process, the respective rationalities of Justice and Prudence together provide the structure from which the utilitarian principle of Universal Good, or Universal Benevolence, can itself be rationally derived. Deontological principles thus crucially endorse the utilitarian principle as a rational, ultimate end, which is also right absolutely, and categorically. Once this principle has been derived, and revealed as self-evident itself, this utilitarian aspect then contextualises and confirms the role of the two deontological principles. In this way, utilitarian principles also endorse the deontological ones. The utilitarian principle cannot (and does not), however, alter, eradicate, or demote or supersede the two prior principles. If this was the case then they would no longer be independently self-evident, and without their self-evident and ultimately rational status, there is no basis for the utilitarian principle in the first place. On the other hand, it is recognised that without the utilitarian principle, Justice and Prudence would remain abstract, and without full meaning. Both moral properties are equally important, as neither by themselves are adequate to produce a full moral theory. It is essential then that the deontological and utilitarian properties within Rational Benevolence retain their own characteristics, while also operating as part of this coherent whole. They both contribute independently *and* depend on each other for full realisation. It is this type of connection that is meant by the phrase ‘mutually informing relationship’ – and the connection that I suggest can be posited directly alongside the mutually *exclusive* state of affairs between deontological and utilitarian positions that has pervaded so much of traditional moral discourse⁵²².

⁵²² In both the introduction and chapter 3, I remarked that there was an extent to which this mutually informing relationship might look something like a Cartesian circle, in that deontology and utilitarianism seem to be defined in terms of each other. It has been said that both types of moral property at once validate *and* depend on each other. But how can something validate something else when it refers to that something else for its *own* validation? This would seem to produce a vicious circle of “reasoning”, in which each part of the whole presupposes each other, and thus neither is really explained or defined at all. But, as I hope I have made clear, the relationship between

Rational Benevolence is the most appropriate and effective representation of that relationship, when Rational Benevolence is understood to be a complete moral theory that embodies the functions of the essential deontological and utilitarian properties of morality. 'Rational' represents the rightness, or the Reason, that is common to both the deontological concept of absolute duty and the utilitarian end of Happiness, as per Sidgwick's earlier interpretation of rightness; 'Benevolence', too, is at once an absolute and categorical moral obligation, and representative of the utilitarian end of general Happiness. The two words are united because both concepts are required in order to create a fully comprehensive moral theory, that effectively includes both deontological and utilitarian properties. In creating this coherency between deontological and utilitarian principles, Rational Benevolence thus accounts for the fact that *both* types of morality appear in our day-to-day moral reasonings. It also reconciles them on the basis that both are established as equally reasonable. In fact, this moral theory *requires* that both are reasonable. The reasonableness of the first two non-consequentialist self-evident principles is necessary in order for the utilitarian principle to be established; the reasonableness of that resulting utilitarian principle is required in order to secure the role of those initial principles.

It is not, then, that there is no difference between the two moral properties. The difference between them is vital, not just for the establishment of Rational Benevolence, but also for the respective values that they represent in the human moral experience. It is rather that there is no longer a divide between the

deontological and utilitarian properties does actually avoid becoming a Cartesian circle, as it is not *completely* circular – at least, not in the way that Rational Benevolence is established, through that process of Sidgwick's philosophical intuitions leading to the utilitarian principle. Those deontological foundations must come first. Having an independent starting point allows the relationship to avoid being a vicious circle - but this might then seem to threaten the argument that both deontological and utilitarian properties have equal status. What I have tried to argue however is that whereas the deontological principles are the vital foundation, they are – as Sidgwick says – too abstract. It is only the ultimate rational end of Benevolence, once it has been derived, that can give them full meaning. Deontological properties *justify* utilitarian ones, utilitarian properties then secure the role of deontological properties within a wider moral theory. This part of the relationship *is*, I argue, circular, as they are independent, and inextricable from each other. Whether we begin with the deontological properties or the utilitarian ones, we will be required to reference the other: Deontological properties require something further for contextualisation, utilitarian properties require a deontological foundation as their 'rational basis'. But it is only this end result that is circular. The theory was built linearly, and this avoids the Cartesian Circle. Also, I argue that the circular relationship is actually part of the cogency of Rational Benevolence. There is strength in the unity between deontological and utilitarian principle, in the fact that one cannot be taken away without the other collapsing. This, on my view, is an effective representation of the permanence of both moral properties in lived moral experience, and of the fact that they are inherently connected.

principles⁵²³. Within Rational Benevolence, deontological and utilitarian properties of morality are made indivisible, both practically, and philosophically. Both are also made complete. Here, we get the holistic view of the two ‘conclusions’ for which Sidgwick argued in the connection passages. The deontological principles would be left inadequate without the utilitarian principle because they require that principle to give them meaningful value; the utilitarian method of ethics would be incomplete because it would be missing the proof of the rational first principle which enables it to provide that support in the first place. And this, I argue, is the truth that lies between those two conclusions.

4.5. Rational Benevolence as a Synthesis

The principal aims of this thesis are to demonstrate that both deontological and utilitarian principles are reasonable aspects of morality, that they are both indispensable to a full understanding of morality, and that they are inherently connected in such a way that they are no longer mutually exclusive, but actually mutually informing. This chapter and the previous two have been engaged with developing this understanding exegetically from Sidgwick’s argument for Rational Benevolence. This section moves away from the exegetical focus, and concentrates now on the capacity of Rational Benevolence to meet those aims. Most specifically, it examines whether Rational Benevolence can meet the given criteria of ‘synthesis’,

⁵²³ Sidgwick himself has already been quoted above as having observed this dissolution of the distance between deontological and utilitarian principles. However, it is conducive here to quote the final pages of *ME*, in which Sidgwick specifically uses the term ‘Rational Benevolence’ for the last time, and in a particular way. Here, Sidgwick is discussing some of the more difficult demands of Benevolence, such as the relief of distress and calamity, sympathy, and possibilities such as that a person may find that they can only promote the general Happiness by working alone, or by grieving those they love, or by working towards ends that they themselves might never see realised (IV:503), ‘In short’, Sidgwick says, ‘there seem to be numberless ways in which the dictates of that Rational Benevolence, which as a Utilitarian he is bound absolutely to obey...’ Sidgwick does not exactly present Rational Benevolence as a complete system of morality here, but he *does* present the case for Rational Benevolence being understood as the crucial informant of utilitarianism, appearing as it does here in its role of providing utilitarianism with the unconditional ought, while also maintaining a certain distinction between that deontological basis of Rational Benevolence and the utilitarian principle itself. (It should also be noted that that at no point in these pages or anywhere else does Sidgwick simply refer to Rational Benevolence as ‘utilitarianism’). It is a perfect example of Sidgwick’s having effectively assimilated the deontological properties of Rational Benevolence with his utilitarianism; it encapsulates the congruence between the deontological basis of Benevolence, and the inherent utilitarianism to which it also points. Sidgwick is right when he says that no-one has taught utilitarianism in this way (III:XIII:387).

as I argue that a synthesis would be the most effective way of incorporating those various understandings of morality into one coherent whole.

‘Synthesis’, it will be recalled, was defined as representing a type of relationship in which the components are inextricably and necessarily linked. I identified three main features of this connection that summarise its purposes: Actuality, recognition, and role. Actuality represents the idea that both moral properties refer to each other for their own full realisation – that in this way, the two properties are fully established, made real, and justified. Because each property is fully realised and justified via reference to the other, both principles must be recognised in their own right (i.e. as having their own identity) in order to continue being an effective informant to the other. This, in turn, leads to the idea that the role of each property for this moral theory can only be understood with reference to the other property. Whichever principle is taken first, it indicates the other. When all these criteria are met, the result is a mutually informing and dynamic synthesis, in which the properties are seen to be inextricably connected because they *require* each other. This requirement also, I suggest, gives that relationship a degree of necessity.

Beginning with the question of whether the deontological and utilitarian moral properties within Rational Benevolence refer to each other for their own full realisation and thus actuality, it is clear that they do. Deontological principles provide the basis from which the utilitarian principle can be derived – but it is the utilitarian principle that gives those deontological principles real applicable value. Each moral property is thus responsible for the establishing and validation of the other. As such however, and as has been seen, this mutual realising can only happen if both moral properties retain their own identities. This is because the deontological and utilitarian properties each represent a different and essential part of the human moral experience. The deontological properties represent the idea of non-consequentialist rightness that does not depend on fitness to an ulterior end, the utilitarian principle represents the idea of Universal Good, or Happiness, as the rationally ultimate end of moral endeavour. But it is that deontological rightness that establishes the utilitarian principle as the rational ultimate end, and the utilitarian principle that confirms the need for those deontological principles as its rational basis. In this way, each moral property depends on the other for their *role* in the construction of the theory of Rational Benevolence. Whichever principle is taken first, it points to its

reliance on the other⁵²⁴. Deontology indicates utilitarianism because the self-evident principles of Justice and Prudence logically progress towards the establishment of Benevolence as the ultimate maxim; utilitarianism indicates deontology because it requires those sorts of principles as its rational basis, or proof.

Rational Benevolence then, I argue, offers a moral theory in which deontological and utilitarian principles are truly indivisible. The properties are inter-reliant, interdependent, and necessarily so. Neither deontological properties or utilitarian ones are fully complete without reference to the other. Even if the two types of property are seen as independently valid, my case for Rational Benevolence argues they still refer to each other in some way, for that very validity. Deontology and utilitarianism are thus engaged in a dynamic relationship which overrides the traditional divide between those two schools of thought. Furthermore, both types of principle are present and necessary on account of their equal *reasonableness*. The deontological properties are (as per Sidgwick's definition of rightness), ultimately reasonable – they then rationally lead to the creation of a rational basis from which the utilitarian principle can be rationally derived as a rational, ultimate end. Not only, then, does this account for the dual reasonableness that I argued in the introduction to be necessary for a fully comprehensive and more accurate understanding of the reality of human morality, but through Sidgwick's understanding of the connection between Reason and *rightness*, both deontological and utilitarian considerations are, in Rational Benevolence, made right on the basis of their respective rationalities. Furthermore, the connection that they share within the synthesis is the direct result of this shared appeal to Reason. It is only on the basis of Sidgwick's particular argument for rightness as similarly applicable to both certain duties and ultimately rational ends that both components in Rational Benevolence can be established as being rationally *right* in the same way. I argue therefore that Rational Benevolence

⁵²⁴ The 'role' aspect of the synthesis also indicates that at a practical level too, absolute rightness and outcome based rightness are linked because each depends to some extent on being informed by the principles of the other. This is essentially where I believe the true normative value of Rational Benevolence to lie. To expand on this, it was said in the introduction that it is easier to conceive of how the dual reasonableness of two different moral approaches would lead to a situation in which they limit each other, and harder to see immediately how they might *inform* each other. I therefore stated that it would mainly be the purpose of the thesis to demonstrate this mutual informing – and it is my view that Rational Benevolence does this effectively. However, it is precisely because of that mutual informing that the limits can also be seen. The scope for mutual validation and mutual limitation between deontological and utilitarian moral properties is examined in the thesis conclusion, where I assess the potential of Rational Benevolence as a theory of normative ethics.

offers a sound and effective synthesis between deontological and utilitarian principles of morality – and, crucially, that this synthesis depends directly on the fact that Rational Benevolence finds *both* types of moral property to be reasonable.

Sidgwick's moral theory could be described as a deontological/utilitarian one, or a utilitarian/deontological one. Neither would be inaccurate. But also would neither fully capture the distinctive assimilation between those principles, in which both deontological and utilitarian principles are rational, self-evident, and *right*, and in which Benevolence has at once a deontological basis, and a utilitarian character. Both of those descriptions are also unavoidably awkward. The only term that satisfactorily embodies and summarises the very particular relationship between deontology and utilitarianism that Sidgwick's theory demonstrates is, quite simply, Rational Benevolence. Only Rational Benevolence embodies the idea that the utilitarian and deontological aspects of morality can be *reconciled* into a coherent whole, that accepts both types of moral reasoning, and is thus reflective of the human moral experience.

4.6. Sidgwick's Synthesis, in His Words: 'Professor Sidgwick's Account of the Development of his Ethical View'

The preceding sections have, I think, adequately summarised the findings of this chapter on Rational Benevolence, and for that reason another conclusion on the premises and mechanisms of Rational Benevolence before these points are stated again in the thesis' main conclusion does not seem strictly necessary. However, there is a final point that must be addressed, and the discussion of this provides a slightly different conclusion on the mutual relationship between deontological and utilitarian principles of morality, that is nonetheless just as useful to my argument. For this reason, the discussion of this final point will serve as the chapter's overall conclusion on Rational Benevolence as a synthesis.

The outstanding issue is that Sidgwick stated in the introduction to *ME* that he would *not* be attempting a synthesis in this volume (I:I:13). He also reiterates in the book's closing chapter that 'a complete synthesis of these different methods is not attempted in the present work' (IV:496). The questions arise therefore as to whether this present work has simply taken liberties with *ME*, and whether the charge could be brought against this theory that to claim that an essential and mutually informing

relationship between deontological and utilitarian principles of morality like the one described above goes beyond the actual intentions, scope, and indeed results of *ME*. It could be asked whether the theory that Rational Benevolence represents an understanding of human morality in which both deontological and utilitarian principles are inherently connected is just conjecture – and/or posited that the theory that Sidgwick perceived the antipathy between deontological and utilitarian principles to have been dissolved has been invented, or forced. Ultimately, it could still, even at this stage, be argued that Sidgwick really was a pure utilitarian, and all that *ME* does is prove that he perceived the utilitarianism represented by Universal Benevolence to in the end be the supreme moral principle, to which all other moral principles are finally found to be subordinate. This last point is particularly important. It is not difficult to support it with the text of *ME*, and even leaving aside the fact that 115 years of history has almost unanimously labelled Sidgwick as a utilitarian, Sidgwick himself did not explicitly dispute that label in his own lifetime.

Fortunately, however, my claims that Sidgwick is only a utilitarian on a *very* specific basis, and that he did perceive and develop this dissolution of the antithesis between Intuitionism (with a capital ‘I’) and utilitarianism, do not rely only my arguments put forwards in this thesis. There is also the support of the personal document, in which Sidgwick gives his own full account of his moral position. Where Sidgwick’s arguments in *ME* sometimes require some piecing together (his argument for utilitarianism, for example, is hardly linear) the PD provides a direct, chronological, and unambiguous account of how he developed his moral theory – and it concurs precisely with the argument that I have given above, for the inherent relation between utilitarianism and deontology. The PD has been discussed in previous chapters, but there it was specifically for Sidgwick’s use of Kant, within the context of arguing that Sidgwick’s Intuitionism shares crucial characteristics with his epistemic intuitionism. Here, it is examined more broadly, for its evidence of the mutually informing relationship overall. The document pulls together Sidgwick’s utilitarianism, his crucial use of a self-evident deontological principle, and the nature of his epistemic intuitionism – and demonstrates the fact that within that intuitionism, there exists a close and necessary connection between his utilitarian and deontological ideas, which Sidgwick did eventually discover and argue for.

Sidgwick begins the account by stating that his ‘first adhesion to a definite Ethical system was to the Utilitarianism of Mill’ (*ME*: xv). This was the relief he had

been wanting from the ‘apparently external and arbitrary pressure of moral rules’, with which he had been brought up. These Intuitionistic rules, he felt, were doubtful, confused, unreasoned – and incoherent. Mill’s system gave a place to the concepts of pleasure and happiness, which Sidgwick thought to be an entirely natural part of human existence. The first problem that Sidgwick encountered with Mill’s system however was that it failed to reconcile self-interest and moral duty. At first, this was the problem with which Sidgwick was the most concerned (*ME*: xvi)⁵²⁵. Left with the choice between placing either the general happiness or one’s own as the ultimate end, Sidgwick perceived that there needed to be grounds on which he could simply see that one should sacrifice one’s own happiness for the sake of the general good. This need, Sidgwick explains, he equated with the need for there to be (despite his initial) aversion to this system - a ‘fundamental ethical intuition’ (*ME*: xvi). The utilitarian system of Mill could not, Sidgwick thought, be made coherent without it.

Sidgwick then describes how ‘in this state of mind’, he returned directly to Kant’s ethics. This is the first of Sidgwick’s recognitions of the possible respective importance of both utilitarian and deontological principles. This time more receptive to Kant’s ideas, Sidgwick describes how he was ‘impressed with the truth and importance of his fundamental principle; ‘*Act from a principle or maxim that you can will to be a Universal law*’. The importance of this fundamental deontological premise – that ‘whatever is right for me must be right for all persons in similar circumstances’ - was, for Sidgwick, the basic intuition of which he had been in search, and was concretely established (*ME*: xvii)⁵²⁶. This maxim, he believed, was both fundamentally true, and practically important.

Sidgwick then describes, in direct accordance with the relevant pages in *ME* (i.e. III:I) that have been discussed above, how Kant’s principle did not seem capable of providing a system of actual, practical duties. At this point, Sidgwick became ‘a disciple on the loose’ - and it is here that Sidgwick can be seen to be standing at the

⁵²⁵ Much of the personal document details Sidgwick’s perception of and attempt to resolve the Dualism of Practical Reason. This issue is immediately detachable from his conclusion on the relationship between utilitarianism and Intuitionism however, which he *did* find to be one of reconciliation.

⁵²⁶ Again, it is possible to draw attention here to the connection that Sidgwick is creating between Kant and intuitionism/Intuitionism, understood *both* as a method of ethics and as Sidgwick’s own form of epistemology. It can also now be seen, in light of my argument for Rational Benevolence, that it doesn’t matter that Sidgwick did not agree with the metaphysical basis of Kant’s maxim. Sidgwick provides his own epistemological justification for it (in the form of his principle of Justice), which he believes to establish its intended objectivity more effectively.

intersection between Intuitionist ethics (which we now know Sidgwick accepts on the basis of Kant's provision of a fundamental intuition), and utilitarianism. As Sidgwick says, '...this led me to reconsider my relation to Intuitionist Ethics. The strength and vehemence of Butler's condemnation of pure Utilitarianism...naturally impressed me much. And I had myself become, as I had to admit to myself, an Intuitionist to a certain extent' (*ME*: xix). He qualifies this with the immensely important and telling statement that 'the supreme rule of aiming at the general happiness, as I had come to see, must rest on a fundamental moral intuition, if I was to recognise it as binding at all'. That fundamental moral intuition, it has been seen, Sidgwick had taken directly from Kant. In this sentence therefore, there is incontestable proof that Sidgwick perceived the utilitarian position to require, as its rational *basis*, support from the Kantian position⁵²⁷. Sidgwick then declares himself have accepted aiming at the general happiness only on the basis that it is itself, an intuitionist premise (*ME*: xix)⁵²⁸. Both utilitarianism and the Kantian maxim appear before Sidgwick here as fundamentally true, and fundamentally important. Sidgwick wondered at this stage whether or not he might have 'a system of moral intuitions' (*ME*: xx). This phrase 'system of moral intuitions', appearing as it does after Sidgwick's recognition that intuitionism pertains as much to utilitarianism as it does to Intuitionist Ethics, clearly represents the fact that within that intuitionism, both utilitarian and Kantian-deontological principles are included.

Sidgwick decides that impartial reflection on current opinion would be the most effective way to test this system. However, that impartial reflection only revealed to Sidgwick the differences between those principles, and his two intuitions of the Kantian maxim, and utilitarianism. This latter principle however, Sidgwick

⁵²⁷ This is, translated into the arguments of *ME*, the stage at which Sidgwick recognises that utilitarianism is in need of proof, and that this can only come from the self-evident principles, one of which had its origins in Kant.

⁵²⁸ Here we see clearly the very close relationship between Sidgwick's own epistemological intuitionism and the Intuitionism that he associates with the Kantian position. There is a clearly still a distinction in Sidgwick's mind, given that he refers first to a 'fundamental moral intuition' and then separately to the 'Kantian principle'. We also have Sidgwick's point that to employ intuitionism means simply 'to see' that something is right – and this is not directly associated with the deontological position, Kantian or otherwise understood. However, Sidgwick's capitalising of 'Intuitionism' in these pages, his association between his need for a fundamental primary intuition and the Kantian maxim on *ME*: xvii, his second and third references to Kant alongside intuitionist ethics (*ME*: xix, xx) and his clear association between Intuitionism and the deontological position that is seen in *ME*, still reinforce my argument that by the time he wrote *ME*, Kant was a vital part of the 'Intuitionism' with which Sidgwick describes the non-consequentialist method of ethics. It also supports my further argument that the lines between Sidgwick's own intuitionism and that Kantian influenced Intuitionism are often indistinct.

states, he found to be ‘in perfect harmony with the Kantian principle’ (*ME*: xx). Where common sense morality itself was not adequate as a system of ethics, and full of contradictions and confusions, Sidgwick’s own two intuitions – one deontological in its non-consequentialist self-evidence, and one fundamentally utilitarian, combined to create a moral theory that Sidgwick *could* endorse, on both deontological and utilitarian grounds. Sidgwick describes how he saw there to be no distance between the Kantian maxim that everyone could will their action to be a universal law, and the utilitarian principle that this action was to be the pursuit of general happiness, or of Benevolence (*ME*: xx). ‘In fact’, says Sidgwick, ‘it was the only law that it was perfectly clear to me that I could thus decisively will, from a universal point of view’. This is a fitting and direct reflection of the argument made above that for Sidgwick, utilitarianism can only be fully understood and validated on the basis of premises from the Intuitionist position. In a very shortened version of the equivalent argument in *ME*, Sidgwick is directly claiming that a version of the deontological Kantian maxim of universalizability and the utilitarian maxim of Universal Benevolence are not only completely compatible, but actually indivisible from each other, within his moral theory. It is at least clear that Sidgwick, despite the decades of argument that he is the most classical of the classical utilitarian, was not prepared to accept utilitarianism at face value: It required, as he describes in *ME*, a rational basis – a rational basis that could only be provided by a maxim that is recognisably Kantian in both origins, and moral-theoretic character. In a poignant and effective summarising statement, Sidgwick declares that ‘I was then a Utilitarian again, but on an Intuitionist basis’.

‘In this state of mind’, Sidgwick continues a few lines later, ‘I published my book: I tried to say what I had found: that the opposition between Utilitarianism and Intuitionism was due to a misunderstanding’. Intuitions are *required* for the basis of Utilitarianism, as we have seen Sidgwick argue in III:XIII. But that Utilitarianism is also completely compatible with those Intuitions, or intuitions. The final words here must be given by Sidgwick, for no other description or interpretation could do them better justice, or better represent the direct correlation between Sidgwick’s own thought on his ethical view, and the argument of this present chapter:

‘I could find no real opposition between Intuitionism and Utilitarianism...the Utilitarianism of Mill and Bentham seemed to me to want a basis: that basis could only be supplied by a fundamental intuition; on the other hand the best examination I could make of the Morality of

Common Sense showed me no clear and self-evident principles except such as were perfectly consistent with Utilitarianism' (*ME*: xxi)

The reciprocal relationship between the deontological and utilitarian properties of morality is thus confirmed by Sidgwick himself. Within this relationship, the two moral properties are interdependent, and inter-validating. They are, essentially, mutually informing.

There remain now only a few comments with which to draw this chapter to a close. They take as their summary the words of E.E. (Emily Elizabeth) Constance Jones, Sidgwick's devotee and one of the people responsible for posthumously organising the material for the *Methods*' 6th edition. E.E. Constance Jones specifically states in her foreword that the personal document manuscript was among the material 'which Professor Sidgwick intended to be referred to, in preparing this edition for the press' (*ME*: xiv). She describes the document as being 'of very exceptional interest', and refers to his passing as 'the calamity of his death'. In both of these judgements, she is not wrong. The suggestion made by the existence of the personal document is that Sidgwick recognised the need to clarify this extremely important topic, and had he lived to engage with it, it is likely that there would have been at least some discourse on this specific issue of the relationship Sidgwick had drawn between deontological and utilitarian premises. From a wider historical perspective, if more interest had been shown in the PD, this may well have altered the world's perception of Sidgwick as being a straightforward utilitarian. He might also have been credited with offering a strong, rational, and highly sophisticated means of understanding how it is that two types of moral principle appear to be reasonable, where most of the history of moral philosophy has traditionally held them to be divided, and opposed. As it was, Sidgwick died on the cusp of those serious changes that were taking place in moral philosophy at the turn of the nineteenth century, and his argument for the relationship between deontological and utilitarian properties – for this *synthesis* - was lost.

THE PROBLEM OF MORAL AMBIVALENCE

REVISITING HENRY SIDGWICK'S THEORY OF RATIONAL BENEVOLENCE AS A BASIS FOR MORAL REASONING, WITH REFERENCE TO PRENATAL ETHICAL DILEMMAS

CONCLUSIONS

5.1. Restatement of the Problem

The thesis' starting point was the observation of the serious moral ambivalence that often occurs in every day moral decision making. I argued that this ambivalence is caused by a conflict between two apparently reasonable and yet fundamentally opposing types of moral principle – that which holds that there are certain moral rules that are absolutely and unconditionally right, and that which holds that it is right to do whatever action will secure the best outcome.

The context of prenatal ethical dilemmas, in which parents are facing a decision regarding the future of a very difficult pregnancy, was used in order to demonstrate this ambivalence. These parents (excluding those who are of a very definitive persuasion either way) are in a position in which both courses of action - the continuation and the termination of the pregnancy – appear in some way to be 'right'. Continuing the pregnancy would be right on the basis that any human life has absolute and unconditional value; ending the pregnancy would be right on the basis that to avoid bringing about painful and distressing outcomes also has unconditional value. What is most crucial for the point about ambivalence is the fact that both courses of action appear to be right, *simultaneously*, despite the fact that they *prima facie* appear to be mutually exclusive. This creates a complicated situation in which each option appears to be morally reasonable, but also morally reprehensible, precisely because of the validity of the alternative course of action. I referred to this situation as 'the Parental Predicament'.

The Parental Predicament also includes the pressure parents often experience from societal attitudes to this issue, which directly reflect the same ambivalence that occurs at the private level. Where individual autonomy is certainly

a vital theoretical framework for protecting parents' decision from external influences, practically speaking it often fails in this remit. Not only does 'freedom of choice' *not* prevent pressure from the public sphere, but it also simply continues to embody and perpetuate any ambivalence. Freedom of choice does not necessarily make those choices easier. I argued in the introduction that for this reason, autonomy itself cannot be the solution to situations of grave moral ambivalence, and that instead the solution would lie in a better and deeper philosophic understanding of the cause of that ambivalence – of the reason why two such different moral impulses exist within our moral experiences, and how both could appear to be right simultaneously. I theorised that this could be understood if the two positions were no longer defended separately on the basis that they are mutually exclusive (which only perpetuates the problem), but rather reconciled in such a way that they can be seen as mutually *informing* – the connection being made by their shared appeal to reason. I argued that if this sort of understanding of morality could be achieved, *both* courses of action in situations such as the Parental Predicament could be better ethically justified, and therefore both more robustly defended at once.

The obstacle to such a solution was identified as being a fundamental problem that exists within moral philosophy (and of which the Parental Predicament is a direct representation): The traditional opposition between two major school of moral thought, deontology and utilitarianism. Representing non-consequentialist absolute principles and consideration of outcomes respectively, the two types of principle could be considered to each describe an equally valid aspect of the human moral experience. Instead, however, a 'divide' between the two types of moral principle sprang up in the early era of modern moral philosophy, with the deontological tradition's emphasis on intrinsic rightness of actions being held at irreconcilable odds with the utilitarian emphasis on the attainment of certain goods, or ends. I argued that this divide – which has continued to pervade much of moral philosophy ever since - creates the artificial idea that morality is exclusively either one type or the other, when actually it is neither completely. Situations such as the Parental Predicament demonstrate that *both* absolute-based and end/outcome-based ethical principles have a role in morality – and that a means of understanding this to be the case is essential for a more accurate, sensitive, and complete comprehension of the human moral experience.

5.2. Restatement of the Solution: Henry Sidgwick's Rational Benevolence

I claimed at the outset of the thesis that Victorian moral philosopher Henry Sidgwick recognises the value, and the role, of both deontological and utilitarian properties of morality in his work. Sidgwick is almost universally considered to be a utilitarian – and often the best utilitarian of the classical era. Whereas I agree that Sidgwick is a utilitarian in some sense, it is my belief that his seminal work *The Methods of Ethics* is not simply a defence of utilitarianism. Sidgwick begins *ME* from observations very similar to mine, which are that all individuals incorporate a range of differing moral principles within their own private moral processes, and that these principles also differ widely between individuals. Pointing out that ‘we cannot regard as valid reasonings that lead to conflicting conclusions’ (I:I:3), Sidgwick’s investigations into these various reasonings eventually lead him to develop a theory in which both deontological and utilitarian principles are necessary for a full understanding of the reality of morality.

To the end of drawing out the process through which Sidgwick arrives at that conclusion, I argued first in chapter 2 that there is a distinct presence of non-consequentialist deontological principles in Sidgwick’s work, and that Sidgwick ascribes far greater value to these principles than is commonly thought. I argued that his understanding of ‘Intuitionism’ (in the moral-theoretic sense) is that it is the non-consequentialist absolute-principle based ethical position, and that his view of moral rightness is that a) it is determined by what is rational, b) it does not involve reference to ulterior results, and c) it takes the form of categorical imperatives. I argued further that Sidgwick’s own moral epistemology actually includes these two largely deontological notions, of non-consequentialism and the connection between rightness, rationality, and ‘ought’. These qualities appear in Sidgwick’s establishing of Justice and Prudence as the only moral maxims that really do admit of being called self-evident. As such, the deontological ideas of non-consequentialist and unconditional moral obligation were found to have a significant role in Sidgwick’s establishing of his last self-evident principle – the utilitarian maxim of Universal Benevolence, which emerges from the combination of Justice and Prudence.

It was then argued in chapter 3 that Sidgwick’s utilitarianism is not always typical of the utilitarian position, in that he emphasises early on in *ME* that the same concepts of ‘right’ and ‘ought’ apply to ultimate rational ends in exactly the same way

as they do to categorical, unconditional duties. Sidgwick argues that rightness does not lie simply in teleological means to attaining ends, but in the rationality of ultimate ends themselves, which only then brings an 'ought' to the means (a situation that is directly reflected by the relationship between utilitarian and deontological properties in Rational Benevolence). In this way, Sidgwick's argument for Happiness as a rational, ultimate end shares much in common with his argument for non-consequentialist actions, or absolute duties. In both cases, his argument for moral obligation is dependent on the shared notion of rightness being that which is ultimately reasonable. It was also found in chapter 3 that Sidgwick's own position on utilitarianism and its dual positive/negative relationship with common sense morality required him, in Book IV of *ME*, to reference the connection between self-evident principles and utilitarianism that he had already developed in Book III. This is, I argued, a reflection of Sidgwick's earlier comment regarding the respective roles of common sense morality and utilitarianism which read, 'the truth lies between these two conclusions. There are certain absolute practical principles...but they are of too abstract a nature....; particular duties have still to be determined by some other method'. Sidgwick was seen to rephrase and emphasise this idea in Book IV with the second of the connection passages, in which Sidgwick suggests that we need a 'line of argument' that allows common sense maxims a certain validity, while also recognising that a further principle is also required to complete them.

The philosophical construction of this connection was then pursued in full in chapter 4, where the thesis' central argument was made that Sidgwick's interpretation of rightness, his subsequent argument for a Kantian influenced self-evident principle, and the logical movement of that principle to the self-evident principle of Prudence, combine to create an argument for impartially realised Universal Good as a rational ultimate end – a moral theory represented by the term 'Rational Benevolence'. In Rational Benevolence, the deontological and utilitarian properties of morality are no longer mutually exclusive, but rather mutually informing, existing together in an interdependent and reciprocal relationship that I have called a 'synthesis'. Sidgwick argues first for the rational, unconditional and non-consequentialist bases of Justice and Prudence, which makes them *right* as per his earlier definition, and gives the resultant Universal Benevolence a true, rational 'ought'. The deontological method through which Benevolence has been derived thus ultimately reveals a utilitarian principle. But that utilitarian principle remains

dependent upon those self-evident deontological principles for the establishment of its own rational basis – that is, for its *proof*. Equally, the self-evident principles rely upon the ultimate utilitarian principle to give them context, and full meaning as part of a system of morality. This shared appeal to the role of Reason, in which the rationality of deontological properties informs the rationality of utilitarian properties and vice versa, reconciles and unites the two types of principle, and is the direct result of Sidgwick's accumulative arguments for reason-based rightness as applying similarly to both absolute principles and ultimate ends, as seen throughout chapters 2 and 3. It is for this reason that I argue Sidgwick's intuitionistic epistemology to be both deontological, (or 'Intuitional') and utilitarian, and that his inclusive moral theory is most conveniently described with the term '*Rational Benevolence*'.

Sidgwick's theory of Rational Benevolence therefore successfully accounts for the equally reasonable basis of the two types of moral principle - absolute principle based and outcome based- that both seem to be so relevant to the lived experience of morality. When Sidgwick remarks that 'we cannot regard as valid reasonings that conflict', he is still right. But, contrary to his prediction that one view or the other would have to be discarded on the basis of its not being reasonable, this was not to be Sidgwick's approach to the problem: Rather, he accepts the rationality of both positions, and *removes the conflict* between them. We need the concept of non-consequentialist morality, and we need it as Kant so famously established it – on a rational basis that can be seen to be self-evident, and thus absolutely applicable. Yet ultimately we find, if we pursue that form of morality, that we must logically come to the utilitarian principle of Benevolence (i.e. to the aiming at Universal Good). Both of these ultimately rational concepts are required for a comprehensive understanding of morality. In Sidgwick's own words, the 'opposition between Utilitarianism and Intuitionism' is 'due to a misunderstanding' (*ME*: xx) – the misunderstanding that there is a fundamental conflict between the ultimately reasonable principle of aiming at the general Happiness, and the traditionally deontological idea of morality being comprised of ultimately reasonable, absolute, and self-evident principles.

Ultimately then, 'we have found that the common antithesis between Intuitionists and Utilitarians must be entirely discarded'. This phrase is, in short, Sidgwick's own abolition of that long-standing divide between deontological and utilitarian positions that has permeated so much of moral philosophy, and that is seen in the lived experience of deep moral ambivalence. Showing that this

ambivalence is the result of the *equal reasonableness* of both deontological and utilitarian moral properties, and providing a theory of how the properties are united within a coherent whole, is, I believe, the most effective means of challenging that divide – and Rational Benevolence can perform both of these roles. I therefore argue that Rational Benevolence provides a reconciliatory solution to the philosophical debate between deontological and utilitarian moral views, and that subsequently it offers a means of understanding and possibly managing moral ambivalence as it is actually experienced.

5.3 Using Rational Benevolence as a Moral Theory

It has been the chief aim of this thesis to demonstrate that the human moral experience is comprised of neither absolute-principle based or outcome based ethical principles exclusively, and that instead both types of moral value are reasonable, and important. I theorise that my argument for the mutually informing synthesis between deontological and utilitarian properties of morality that is embodied by Rational Benevolence mounts a persuasive theoretical case for this view, and thus the thesis' main objective has been met. But this argument for Rational Benevolence has, throughout the thesis, been made entirely in conceptual terms. In order to truly assess its efficacy, it remains to consider the extent and range to which Rational Benevolence can function as an applied moral theory, that might be capable of giving actual moral guidance.

I stated in the introduction that it is beyond the scope of this thesis – and certainly of its conclusion – to adequately expound and discuss all the details of how Rational Benevolence might apply in practice, as a theory of normative ethics. It is however possible, I believe, to at least give some rudimentary outlines of how Rational Benevolence seems to qualify as a functional theory. These can be presented in two over-arching observations about the theory, that can then be laid down as the main lines of enquiry regarding how Rational Benevolence might be able to offer something of practical substance.

Firstly, there is the fact that Rational Benevolence argues comprehensively that *both* deontological and utilitarian moral properties are *right* because *both* are *rational*, and that neither type is expendable or in error (when viewed through Sidgwick's own philosophical arguments) – a characteristic that we can call, for ease

of reference, 'dual rationality'. This statement has what I will call immediate *descriptive* value, by which it is meant that this dual rationality is effective in describing and explaining the reason for moral ambivalence. Furthermore, this dual rationality indicates that the relationship is a dynamic affair, wherein both the deontological and utilitarian properties must retain those respective rationalities in any given situation in order for that coherent whole of Rational Benevolence which makes up 'morality' to be maintained. Given that this is a dictate that must be observed when making moral decisions according to Rational Benevolence, it is this aspect of the synthesis that gives Rational Benevolence a degree of *normative* value.

The second line of enquiry involves recognising that Rational Benevolence is predominantly a theory of rightness, but that it still maintains a secure place for the *good*; the Happiness that Sidgwick considers to be the end of the utilitarian principle is, after all, specifically identified by Sidgwick to be that 'good' (III:XIII:388 ; III:XIV:400, 406-407). The Universal Benevolence of Rational Benevolence therefore begs the question (as is the case with any theory that includes the 'good') of what exactly that good, or Happiness, is – of what it is meant to really *look like* in real life, within the normative framework that is established by dual rationality.

The remaining sections of this chapter, and of this thesis, are engaged in exploring further how the dual rationality aspect of Rational Benevolence, and its preservation of the good, contribute to these descriptive and normative functions of Rational Benevolence. I return to the Parental Predicament for the purposes of demonstration and application, and within this context I aim to explore the possibilities, and the limitations, of Rational Benevolence for practical ethics. Whereas I have only tentatively posited some of the ideas here, I have indicated where further work in this area might lie.

5.3.a. The Descriptive Value of Rational Benevolence

The descriptive value of Rational Benevolence is quite simply that this theory is able to more comprehensively explain at a metaphysical level the complex phenomenon that is the human moral experience, than theories that claim that morality is exclusively either deontological or utilitarian. As I have argued, theories that support one viewpoint or another are likely to always be required to defend themselves

against the alternatives. Ironically, this often causes an increase in the conflict between them, which in turn leads to the perpetuation of that very need for each theory to defend itself. Thus, deontological and utilitarian understandings of morality are often pushed further, and unhelpfully, apart. Rational Benevolence, on the other hand, argues not from one position or the other, but precisely from what I believe to be the common ground between them, where properties of both combine to create a full understanding of what can be called morally reasonable.

In order to understand how this is helpful at a practical level, the deontological and utilitarian properties of Rational Benevolence require being phrased in terms that correspond directly with the ambivalence experienced by parents in the Parental Predicament, as follows.

The deontological aspect of Rational Benevolence - that which upholds the concept of absolute, inviolable, and unconditional principles - typically represents and supports the strong disposition parents are likely to feel towards wanting to continue the pregnancy, regardless of any difficult circumstances. This desire may occur for any number of reasons, such as the intense emotional awareness that this is *their* child, some religious tendencies, or indeed a combination of any or all of them – but the underlying theme that characterises this position is that this pregnancy has an unconditional value, and the continuation of it is therefore unconditionally right. The utilitarian aspect of Rational Benevolence – which supports considerations of quality of life and the potential for flourishing - typically represents and supports the strong disposition parents are also likely to feel towards wanting the best *kind* of life for their child⁵²⁹. This concern for quality of life, and freedom from pain and suffering, is also felt to be unconditional. Ambivalence itself is, of course, caused by the apparent value of both views and the obvious tension between them. The descriptive function of Rational Benevolence first applies when it explains – and justifies - this ambivalence by giving greater philosophical support to the argument that *both* areas

⁵²⁹ I made the point in the thesis introduction that accounts of the two positions such as these I have just given are highly simplified, but I feel it is important to emphasise here also. This simplification is especially true of the utilitarian aspect. Other outcomes that parents are likely to have to consider include the related sense that they are responsible for securing a quality of life for their future child, and that this is only compassionate – and also that acting compassionately is itself difficult to define. For example, they may also have other lives towards which they also need to act compassionately, such as other children, any other sort of dependents for whom they may be carers, and indeed their own lives, especially if the physical or mental health of either parent is in question. All of these factors however contribute legitimately to what it means to wanting to secure the best *kind* of life, for any or all of the parties involved.

of concern, and both courses of action, are indeed as equally valid as they appear. Ambivalence is directly reflecting a major aspect of moral reality, which is that morality *is* comprised of this range of considerations, and that both the deontological and utilitarian properties *can* claim to be reasonable, and therefore morally right⁵³⁰. It is precisely this situation that is embodied by the dual rationality of Rational Benevolence – by its preservation of both deontological and utilitarian considerations as ultimately rational, and therefore ultimately moral.

It might be posited here however that Rational Benevolence still actually offers nothing substantially different to other moral theories that defend either course of action separately, especially given that many of these theories also utilise rationality as a basis for their moral reasoning. By simply claiming that both courses of action are reasonable in their own separate and independent ways, this thesis would contribute nothing towards reconciling the moral philosophical rift between these apparently fundamentally different positions, and would certainly go no way towards creating an understanding of ambivalence that could be used to aid the private and public pressure that it causes⁵³¹. I maintain against such objections

⁵³⁰ In this way, Rational Benevolence can also be used to defend the non-ambivalent courses of action (termination or continuation), because both absolute-principle based and outcome based moral principles are defended independently by Rational Benevolence, as well as together. However, there are some important caveats to this claim. Rational Benevolence, it will be seen shortly, does *not* simply justify either course of action, simply because it ‘covers’ both absolute and outcome based approaches to a situation. Both of those approaches may well be capable of being proven Benevolent, but the whole point of Rational Benevolence is that that Benevolence, and any actions that stem from it, are also *rational*. Thus, limits will start to appear where actions cannot be called rational. This very point is accounted for in the discussion of Rational Benevolence’s normative value.

⁵³¹ This would also offer no aid to the enduring problem in this context of how to define what is meant by concepts such as the absolute value of life, the good life, suffering, or the avoidance of suffering. Who is to decide, ultimately, what constitutes suffering, or ‘a good life’? These questions, and related ones such as whether or not we have a responsibility towards future children now that certain technologies are available, are currently occupying many areas of medical ethics, and the area of prenatal testing especially. For excellent discussion of the issues, see Jonathon Glover, *What Sort of People Should There Be?* (London: Penguin Books Ltd:1984), and *Choosing Children: Genes, Disability, and Design* (Oxford: Oxford University Press: 2008); Michael Sandel, *The Case Against Perfection: Ethics in the Age of Genetic Engineering* (Cambridge MA: Harvard University Press: 2007); David Galton, *Eugenics; The Future of Human Life in the 21st Century* (London: Abacus: 2001); Phillip Kitcher, *The Lives to Come; The Genetic Revolution and Human Possibilities* (London: Penguin Books: 1997); John Bryant and John Searle, *Life in Our Hands: A Christian Perspective on Genetics and Cloning* (Leicester: Inter-Varsity Press: 2004); John Harris, *Enhancing Evolution; The Ethical Case for Making Better People* (Princeton: Princeton University Press: 2007); Tom Shakespeare and Ann Kerr, *Genetic Politics: From Eugenics to Genome* (Cheltenham: New Clarion Press: 2002) (other writers and theorists who are also relevant to this area are listed below, within the more specific context of Transhumanism). It is beyond the scope of this current work to examine this aspect of the debate in detail – however, I do argue below in the discussion on the normative potential of Rational Benevolence, that this thesis may offer indications as to how these ideas might be defined.

however that Sidgwick's Rational Benevolence attributes rightness to both deontological and utilitarian properties of morality more effectively than other theories precisely because Rational Benevolence does not simply defend the reasonableness of each separately. Rather, it embodies the fact that both types of moral property are reasonable, at the same time, within *one* united theory. This then leads to a profounder explanatory advantage of Rational Benevolence, which is that it can explain not just the fact that both courses of action *are* morally right simultaneously, but *why* this is the case. The level of descriptive capacity I have already outlined in this section can be referred to as Rational Benevolence's *superficial* descriptive value, as it explains at a surface level that fact that both courses of action *are* morally right. But this second, deeper, explanatory function of the theory – its ability to explain *why* this is the case – I will call Rational Benevolence's *substantive* descriptive value. This substantive descriptive value lies specifically in the synthesis between utilitarian and deontological moral properties, in that it represents the concept that within dual rationality, the rationality of each position is inherently *connected* to the rationality of the other. The substantive descriptive level signifies the dynamic nature of the mutually informing synthesis between rational deontological and utilitarian principles – which states that both the deontological and utilitarian properties of Rational Benevolence must retain those properties, *and* refer to the other, in order for the full definition of what is morally right to be achieved. And it is in this way that the substantive descriptive value of Rational Benevolence then begins to reveal the normative capacity of the theory.

5.3.b. The Normative Value of Rational Benevolence

It has just been stated that the substantive descriptive value of Rational Benevolence leads directly to the potential normative application of the theory, given that the mutually informing relationship embodied therein implies that a process must occur in which absolute-principle and outcome-based considerations are actively maintained while also demonstrating necessary reference to each other - and this in turn could apply within a given set of ethical circumstances. I theorise here that this normative function of Rational Benevolence could apply both positively, in that it could better validate and justify decisions, and negatively, in that it could offer some

means of establishing appropriate limits on certain decisions⁵³². The validation of action aspect will be discussed first, as it is through the process of validation that the matter of limitation to action is naturally raised. Again, I concede from the outset of this stage of the investigation that this exploration into the normative value of Rational Benevolence is only a cursory one - but the ideas are outlined as extensively as is possible in the space available here.

5.3.b.i. Rational Benevolence as Validation of Moral Action

It must first be asked of this claim made by the substantive descriptive value of Rational Benevolence that the rationality of each moral property informs the rationality of the other in order to make both types right, exactly what it means for two positions to be simultaneously right, or rational, when the practical level involves two fundamentally different courses of action. Two interpretations can be eliminated immediately. Firstly, it does not mean that deontological and utilitarian properties of morality are simply the same. The construction of Rational Benevolence at the theoretical level has shown that this actually must *not* be the case - for the validity of each moral property depends on the properties of the other - and to say that there is no difference between the principles would of course also fail to address the very problem of moral ambivalence with which we started, which is precisely the existence of two very different and yet apparently valid moral impulses. Secondly, it is also not meant by 'simultaneous rightness' that *both* courses of action are an 'ought' – something that could be implied by the role of 'ought' in Rational Benevolence. As Sidgwick's Reason-based definition of right generates the notions of ought and moral obligation in the way that it does, saying that a course of action

⁵³² It will be recalled from the introduction I argued that equal reasonableness must point both to the validation and the limitation of each position. I also stated there that it is easier to establish how two reasonable positions might *limit* each other with their respective reasonable concerns, as opposed to how they might validate each other – and that this thesis would be engaged with building the case for the latter relationship. Now that this has been done, it will be possible to see in these last sections that the mutual validation aspect of the deontological/utilitarian relationship is, in practice, actually *easier* to defend than the mutual limitation aspect. This is especially the case in the context of the Parental Predicament, where establishing *limits* to action could perhaps require either dictating that a pregnancy *ought* to be continued, or that it *ought* to be ended. The first of these two positions can only really be defended tenuously at best, and the second barely at all. This, as I will draw attention to in the main text below, is admittedly a challenge to the normative value of Rational Benevolence. On the other hand, the potential of Rational Benevolence to *justify* certain courses of action, and within *certain parameters* that at least indicate where limits should lie, even if they are not enforced, does seem to be apparent.

can be justified by this understanding of rightness might seem to make that course of action appear an inviolable moral obligation; it is, according to Sidgwick, only rational – and *right* - to pursue that which we recognise as rational, and right. Not only is this practically paradoxical (it logically cannot be the case that we ‘ought’, on a direct understanding of that term, to pursue two entirely different courses of action at once), but it would also lead to many morally unacceptable situations in which it would seem as though Rational Benevolence is dictating that parents *ought* to continue a pregnancy – or, even more deplorably, that they *ought* to end it.

But to give dictates in this way is not how the ‘ought’ in Rational Benevolence is intended to operate. The ‘ought’ in this theory is the crucial component of what establishes the *Benevolence itself* as a moral duty. That is, it is the moral obligation (or the duty) of Benevolence (which prescribes aiming at the general Happiness) that is established on the basis of the ought; Rational Benevolence dictates that we ‘ought’ to obey the duty of Benevolence. This Benevolence is then, as seen in chapter 4, made up of the peculiar balance of deontological and utilitarian properties which allows two different courses of action to be right, or rational, or *benevolent*, according to the evidence given by the construction of Rational Benevolence that morality exists in both. Benevolence as a moral obligation therefore refers to a class of actions, rather than to specific actions themselves⁵³³.

Attaching the ‘ought’ to the Benevolence that gives actions their moral quality leads directly back to the normative question of what exactly these benevolent actions are to look like. As Sidgwick said, we ultimately require of a system of ethics that it tells us what it is actually right to *do*. More precisely for the context of Rational Benevolence, we require to know how benevolent actions embody the crucial dual

⁵³³ Sidgwick himself appears to directly support this interpretation of the relationship between the ought that is derived from Reason, and *classes* of actions, rather than individual actions. At I:III:33-34, during his seminal discussion of ‘ought’, Sidgwick points out that by referring judgements of ought to ‘Reason’, he does not mean to imply that every such judgement is derived from a universal axiom. In fact, he says, ‘the moral faculty deals primarily with individual cases as they arise, applying directly to each case the general notion of duty, and deciding intuitively what ought to be done by this person in these particular circumstances. Sidgwick goes on to explain that he has used ‘Reason’ to denote moral cognition as distinct from mere feelings, but then adds, ‘as a further justification of this use’, that when we have passed moral judgement on a particular case, we generally ‘regard it as applicable to any other action belonging to a certain definable class: so that the moral truth apprehended is implicitly conceived to be intrinsically universal’ (I:III:34). This is precisely how the relationship between Benevolence and ‘ought’ is envisaged to work. An individual would pass moral judgement on a particular case, according to the circumstances of that case, but this moral judgement would be seen as belonging to the class of actions defined as benevolent, which is itself universal – and which *does* incur the notions of ‘ought’ and ‘moral obligation’. The particular action is not an ought, but the class to which it belongs is.

rationality of both deontological and utilitarian moral properties, that defines them as benevolent in the first place. For *this* is the answer to what it means for two different courses of action to be right simultaneously⁵³⁴. Dual rationality directly creates *dual validation* via the fact that each principle retains its own inherent reasonableness within the synthesis. It is only the deontological properties that can reveal the utilitarian principle; it is only the utilitarian properties that then make sense of the deontological principles. Because of this, both types of moral property combine to actively and respectively validate to a *reasonable degree*. This means, in theory, that even when a particular course of action is followed, the rational moral properties on which it is ethically justified (whether they are absolute-principle or outcome based) will still be informed by the rational moral properties of the other type of principle. This balance thus avoids a simple oscillation between the two positions, and instead indicates that an equilibrium between the reasonableness of each principle could be established⁵³⁵. Such an equilibrium would signify a morally sound action i.e. a 'rationally benevolent' action.

⁵³⁴ In the thesis introduction, R.M. Hare's argument that two oughts cannot rationally be oughts at once was examined for how Hare approaches moral conflict, and for how this compares to how Rational Benevolence approaches it. I argued there that both 'oughts' are in fact viable, and that in order to argue this, it needed to be shown that this is not actually irrational. This argument I made on the basis that the dual reasonableness of the deontological and utilitarian properties that make up the Benevolence that is at the heart of 'rationally benevolent actions', and thus rationally benevolent actions can be found in either course of action, provided they embody that mutually informing reasonableness. The argument I am making here is the extended version of that argument, with my having added the qualification that those viable 'oughts' are not to be seen as specific dictates, but rather as embodying the 'ought' of Benevolence, from which they are derived.

⁵³⁵ I do not mean the word 'equilibrium' here to have connotations with reflective equilibrium; it is meant only to represent a meeting point, or a balance, between two positions. However, I do agree that when its practical workings are phrased in this way, Rational Benevolence as a method of ethics does bear more than a passing resemblance to reflective equilibrium, as was discussed in the literature review. If reflective equilibrium is understood in a very general sense to refer to a process through which coherence is sought between a variety of our judgements and beliefs, then Rational Benevolence's normative process of reaching a balance between the various absolute principles and outcome-based values qualifies it as a form of that method. However, one of the criticisms that is often advanced against reflective equilibrium is that it presumes the non-revisability of certain points, and therefore affords these points immediate justification – and because the self-evident deontological and utilitarian components through which Rational Benevolence is constructed could also be seen as non-revisable, this actually makes Rational Benevolence appear *less* coherentist than some forms of reflective equilibrium. On this basis, Rational Benevolence appears to emerge more along the lines of the sort of semi-foundationalist reflective equilibrium, argued for by McMahan (see Jeff McMahan, 'Moral Intuition', in, Hugh LaFollete (ed.) *The Blackwell Guide to Ethical Theory* (Oxford: Blackwell Publishers, 2000) pp.92-110). There is also the fact that foundationalist approaches to morality tend to emphasise the role of epistemically derived *a priori* intuitions, and this is certainly a crucial part of Sidgwick's own epistemology. It would seem then as though Rational Benevolence can offer no further conclusions on whether or not Sidgwick is a coherentist or a foundationalist, or whether or not he is using a form of reflective equilibrium. As was seen in the literature review, there are simply too persuasive grounds for both views.

5.3.b.ii. 'Rationally Benevolent' Actions

In order to examine the specifics of 'rationally benevolent' actions, the actively normative aspect of dual rationality can now be stated in the contextualising terms of the Parental Predicament. Introductory scenario A, in which parents A1 and A2 have received a positive result from a test informing them that their child will be severely disabled, is revisited here in order to give an example, as follows. It is rational to think of a potential/future human life as having unconditional value, and it is equally rational to believe that that life should be free from pain and suffering. But where it might be more obviously called 'deontological' to continue the pregnancy, Rational Benevolence states that that ultimately self-evident, rational, and non-consequentialist principles which sustain the deontological property are what reveal the ultimately rational utilitarian principle of Universal Benevolence, or *Happiness*. That is, the action of continuing the pregnancy on apparently deontological grounds is only rationally justified in the way that it is, because the end of Happiness is rationally embedded within those absolute principles. In this way, the rationality of life's unconditional value *and* the rationality of freedom from pain are both accounted for, to a reasonable extent. Continuing the pregnancy, as parents A1 do, would therefore emerge as a rationally benevolent action, according to the way in which that Benevolence is vitally constructed of interdependent deontological and utilitarian properties. Equally, where it might be more obviously called 'utilitarian' to avoid pain and suffering, Rational Benevolence states that utilitarian Benevolence can only be established according to those prior non-consequentialist, absolute, and unconditional principles. That is, the action of ending the pregnancy on apparently utilitarian grounds is only rationally justified in the way that it is, because those absolute principles are embedded within the establishment of the end of Happiness. Again, the rationality of life's unconditional value *and* the rationality of freedom from pain are both accounted for, to a reasonable extent. Ending the pregnancy, as parents A2 do, would also therefore emerge as a rationally benevolent action, on the same grounds as A1, which is that Benevolence is vitally constructed of interdependent deontological and utilitarian properties.

This is a perfunctory overview of how the two types of moral value inform each other at a practical level; of course, situations will in reality be far more complicated than this formula suggests. But I do posit that the formula itself is viable. According

to Rational Benevolence, both courses of action, whether apparently ‘deontological’ or ‘utilitarian’ at face value, are only made right by the fact that they also refer, for that rightness, to the values of the alternative moral property – and this is because Happiness, as the ultimate rational end, necessarily includes both absolute and outcome based values in a way that they are *all* unconditional. In these decisions, both properties maintain their own identities while also referring to the other for their own validation, and each course of action has *both* the deontological and utilitarian properties that together make up the moral rightness of Benevolence. This is, I argue, a far more robust explanation for the ending and for the continuation of difficult pregnancies, and its strength depends on the fact that it does not discount one set of values, but incorporates them both as essential contributors to moral reasonableness, or rightness.

5.3.b.iii. Rationally Benevolent Actions and Universalizability

There is also a further possible test for the rational benevolence of actions, which arises from the fact that the ultimate moral duty was originally called by Sidgwick ‘Universal Benevolence’, and on the basis that Sidgwick established Universal Benevolence on an adapted Kantian universalisation maxim. This is that rationally benevolent actions – courses of action that have been decided upon via the process of balancing the deontological and utilitarian properties that was outlined above – should, being ultimately rational, technically admit of being universalised. At the normative level of how Rational Benevolence provides validation to action however, this issue of universalizability is admittedly problematic. On one hand, it could be argued that what is being universalised is the *Benevolence* itself. I actually do maintain this, and suggest that it is because it is a *class of actions* that is being universalised that both courses of action can be ‘universalised’ at once, provided the course of action is informed by the reasonableness of both types of moral property and is rationally benevolent⁵³⁶. However, the concept of universalizability in practice

⁵³⁶ I suggest that the universalisation of Benevolence be established on the following grounds. As has been seen in the thesis, Sidgwick underpins his theory of Rational Benevolence with a Kantian-esque maxim of universalizability, adapted by Sidgwick on the grounds that Kant’s rule – ‘what I judge to be right must...be judged to be so by all rational beings’ (III:I:208), although indispensable to any system of rational morality, is not, in *that* form, capable of producing a complete system of moral duties. On Sidgwick’s reasoning, the Kantian maxim must be rephrased as the principle of Justice, and this – combined with Prudence – eventually yields the self-evident, rational principle of Benevolence (each

seems to demand that each individual course of action prove itself to be universalisable, in order to prove itself rational, and therefore right – and this seems untenable⁵³⁷. We would have to ask whether parents who decided to terminate on the basis of cerebral palsy, for example, could reasonably will that all people in their situation do the same, which would incur the idea that all people *ought* to do the same. Whereas it has been argued that this is a viable belief, for most people this is likely to still appear to carry a significant amount of what bioconservative writer Leon Kass called within bioethics, ‘the wisdom of repugnance’, or, in short, ‘the yuck factor’⁵³⁸. It simply does not *feel* ethical that such a rule should exist. It certainly does not accord with the values of a politically democratic and liberal society that only one course of action in such situations be regarded as the correct one.

The problems with universalizability lead into the territory of the long-running philosophical debate over the identity within ethics of ‘the good’. This is because there is within Benevolence – which dictates that actions are to aim at the general Happiness that Sidgwick also equates with Ultimate Good – an inherent flexibility as to how to define that Happiness, or good. Nowhere is this more evident than in the context of the Parental Predicament, where the range of opinions on what constitutes Happiness vary enormously. Where one family may not consider cerebral palsy, or Down’s Syndrome, or blindness (or any condition, in fact) to be a major obstacle to the Happiness of any concerned interests, another family may consider these conditions to be catastrophic⁵³⁹. This clearly presents a firm challenge to the idea that any sort of action under such circumstances could be universalised.

one is morally bound to regard the good of any other individual as much as his own). But that principle of universal and equal Justice is a crucial part of the process from which Benevolence has been derived, and once Sidgwick has identified Benevolence it retains that universalising contribution from Justice. So when the unconditional duty of Benevolence thus demands that ‘benevolent’ actions pass the universalisation test, (understood loosely as ‘whatever is right for me must be right for all persons in similar circumstances’ (*ME*: xvii), or in more specific Sidgwickian terms as ‘it cannot be right for A to treat B in a manner in which it would be wrong for B to treat A’), it is the production of Universal Good/Happiness that is being tested – and *that* which is being universalised. Again, then, it is the maxim of Benevolence, and the class of actions therein, that is capable of being universalised – as opposed to individual actions. (There is, perhaps, a direct line of comparison here between this idea of universalisable but not independently categorical duties, and Ross’s argument for *prima facie* duties).

⁵³⁷ Further to my point at fn.8, Sidgwick’s version of universalizability does not at the theoretical level generate specific unconditional and inviolable duties in the same way as Kant’s version of the categorical imperative, which is why it is difficult to apply it to individual cases.

⁵³⁸ See Leon Kass, ‘The Wisdom of Repugnance’ (in, *The New Republic*, Vol.216, No.2 (1997) 17-26).

⁵³⁹ The situation is further complicated by other factors such as the current inability of pre-natal medicine to detect the actual severity of conditions such as Down’s Syndrome.

But equally, nowhere is this flexible notion of Happiness actually more *important* than it is in this context. It is here that the second line of enquiry into how Rational Benevolence could apply as a practical moral theory – that which identified that a significant place for the good is maintained within Sidgwick’s theory of the right – starts to become both relevant, and indispensable. For no two cases in the area of the Parental Predicament will ever be the same, and it is vital that this diversity of opinion regarding the good is accounted for, particularly within this most sensitive field where the autonomy that should allow parents to make their own decisions as to what is right for them and the happiness of their own family, so often comes under such threat. The fact that Rational Benevolence, although held up as objectively *right*, can also preserve a place for a degree of subjective goodness means that there are few restrictions on exactly what benevolent (or rationally benevolent) actions are to look like. It therefore actually works to the enormous advantage of Rational Benevolence that it allows for this varied interpretation of goodness, or Happiness, despite this posing a challenge to the applicability of universalizability.

On the other hand, universalizability *does* maintain a certain vital check upon those interpretations of ‘happiness’ and ‘good’. This further delineates the role of the good in Sidgwick’s theory, and also lends universalizability a greater degree of potential to be incorporated into Rational Benevolence as normative theory. This is, once again, the direct result of the *rational* basis of Benevolence, and the synthesis between absolute-principle based and outcome-based concerns that Rational Benevolence embodies. As has been said above, this synthesis depends on the mutual rationality of both types of moral property. This immediately indicates that if either property was to lose its reasonableness, the decision in question would lose its moral justification, and a *limit* would appear to that particular action. In other words, a limit would appear on what can truly be called Benevolent, or *good* – and Universalizability could well be an effective measure of this. It is to the idea of Rational Benevolence posing *limits* to moral action that the next section now turns.

On the other hand, there are other extreme examples of the enormous range of ideas about what constitutes a good life, such as the community in Maryland, Baltimore in the USA that has been affected by profound deafness, and has simply developed their societal norms around it. Within their own community, individuals are not ‘disabled’ as other hearing communities would consider them to be, and deafness is simply a part of their individual and communal identity, and their ability to belong. On this basis, some parents in this community who have had the opportunity to do so have specifically selected *for* deafness.

5.3.b.iv. Rational Benevolence as Limitation to Moral Action

Thus far, the discussion of Rational Benevolence as a theory of normative ethics has only included the positive angle, of how the mutual reasonableness of both deontological and utilitarian properties combine to validate and support courses of 'rationally benevolent' action. But there is an immediate converse to this relationship in which two types of property have equal worth, which is the recognition that if each property is going to retain its own independent validity and rationality in order to sustain and validate the other, then there must come a point at which this validity and rationality must impress limits on the validity of the other. Practically speaking, if each property can validate certain courses of action to a reasonable degree, then by its own reasonableness it must logically be able to limit certain courses of action, also to a reasonable degree. If either value (absolute-principle based or out-come based) ceases to be present, or rational, this will become apparent by the fact that the two properties that make up rationally benevolent actions cannot be seen to be mutually supportive in the decision. If this is the case, then the action in question will not be rationally benevolent – that is, it will not be *right*. In the context of prenatal medicine in particular, examples of this are not difficult to find. One such example might be some terrible scenario – theoretically possible – in which the parent(s) intend to continue the pregnancy, despite the fact that the welfare of numerous parties with interests connected to the pregnancy (the foetus, the mother, the father, any other dependents) would be at serious risk; an example from the other course of action could be where parents are determined to end a pregnancy on ethically weak grounds, such as sex, or even eye colour⁵⁴⁰. In the first case, it seems difficult to maintain that this decision represents a process in which absolute-principle based

⁵⁴⁰ With the advent of more and more technologies that are designed to test for traits such as eye and hair colour, and even complexion, screening and selection for these traits – and the implications of non-selection – are on a steady rise. In relation to this screening for non-medical conditions, assisted reproduction companies such as Fertility Institutes in Los Angeles have implemented the phrase 'severe cosmetic conditions', and use of these services is growing, both in America, and in the UK. It should be noted that this practice is attached to embryo screening, as opposed to natural pregnancies, and therefore a 'termination' is of a different nature, and according to many theorists – and users - not as ethically complex. But whereas pre-implantation genetic diagnosis has been in many areas an important tool in the avoidance of certain genetic conditions, it still brings with it its own set of ethical concerns, especially now that use of screening has extended to the non-medical – and both moral philosophy and applied ethics must keep abreast of the issues. I speculate here that Rational Benevolence could be applicable to the implementation of these technologies, although I am aware that in this particular field, it may be too late.

and outcome-based principles have arrived at an equilibrium, and it appears as though the balance has gone in favour of absolute-principle considerations only. The equally crucial outcome based aspect of Benevolence has been disregarded, and the decision does not therefore emerge as rationally benevolent. In the eye-colour scenario, it would seem that the absolute value of human life has been disregarded, on grounds that have no moral weight— and again, in the absence of this value, a decision to terminate on these grounds cannot be called *rationally* benevolent⁵⁴¹.

These two examples have been purposely chosen to represent the extremes of possible decision making in this context – because it is in reference to examples such as these that the universalising aspect of Rational Benevolence might be more applicable. Universalizability struggled in terms of *validation* to action, because it was seen that it would be very difficult in this particular context to justify a course of action on the basis that ‘everyone should do x’. In terms of *limitation* to action however, i.e. in dictating what ought *not* to be done, universalizability appears to have a more legitimate role. This is again due to the flexible nature of the ‘good’. Where it was seen above that it is to Rational Benevolence’s advantage that it protects a varied interpretation of what is good, or Happiness, we know that not *all* interpretations of good are truly benevolent, because not all interpretations of good are *rationa*⁵⁴². Rational Benevolence suggests here that if a ‘benevolent’ action truly cannot be universalised, then it is not actually rational, and therefore not actually benevolent, and therefore neither morally *right*, nor good. This holds in the case of the eye colour scenario. However much the parents may believe that Happiness is only achievable if one has blue eyes, this hardly seems a rational belief. And it does not seem rational to universalise that all pregnancies bearing future people with

⁵⁴¹ It could be imagined that the parents *genuinely* believe that life is not worth living unless one has blue eyes, and that therefore according to them ending the pregnancy is a genuinely benevolent action. But a genuine belief that life is not worth living unless one has blue eyes does not seem to be a rational one. It might, in fact, indicate serious issues with the mental health of the parent(s). In this case, termination of the pregnancy would probably become rationally benevolent on the basis that the parents are not well enough to cope with the pregnancy/baby, and the possibility that the future child may not be adequately cared for. In this case however, the ending of the pregnancy has been justified i.e. has been made rationally benevolent, on entirely different grounds from that of eye colour, and eye colour remains an irrational and non-rationally-benevolent reason for termination.

⁵⁴² The most obvious example of the socio-political damage that can be wreaked when ‘good’ is held up to the extent to which the concept of objective right – or the human rights therein – is almost completely obliterated is that of Nazi Germany. The Third Reich’s systematic and devastating attack on the Jews was presented, to the public at least, as necessary for the building of a greater, safer, *better* Germany, complete with a *better* population. With this posited as an end most important for the ‘good’ of the country, mass genocide simply became a means to that end, with little or no significance of its own.

brown eyes be ended; we could also plausibly universalise the alternative, and say that such pregnancies *reasonably ought* to be continued⁵⁴³.

The case of the parents who decide to continue despite the serious risks to health, however, is a different matter. On the one hand, the most obvious argument from Rational Benevolence is that as the child will be born into nothing but pain and suffering and will have no chance of flourishing, then the duty of aiming at general Happiness seems to have been violated, and this would appear to not be a rationally benevolent action. Certainly, this course of action would not seem to be universalisable, especially according to Sidgwick's maxim that 'it cannot be right for A to treat B in a way in which it would be wrong for B to treat A'; as the parents would presumably not want to experience such pain and suffering themselves, then it does not seem reasonable that they would allow the future child to experience it⁵⁴⁴.

But universalisation has once more run into the challenge that it faced in validation to action, which is that it is very difficult to attach a true 'ought' to that course of action. That is, it is very difficult, in *this* context especially, to dictate what ought to count as 'good', or Happiness, and for as long as this remains the case, universalizability will only have a limited/restricted role⁵⁴⁵. As for the ability of

⁵⁴³ The grounds against eye-colour selection, for example, include arguments such as a) this would lead to a devastating number of terminations, b) it would create a 'gateway' situation in which the lines as to which conditions can and cannot be selected for become so blurred as to fail to be able to implement any policy at all, and, c) that there could quite possibly be a long-term detrimental effect on the gene pool, which would now be lacking in the genetic diversity that is crucial for human health, and survival.

⁵⁴⁴ The rest of the principle of Justice reads 'without there being any difference between the natures or circumstances of the two which can be stated as a reasonable ground for the difference of treatment' (III:XIII:380), and when this caveat is added, it can be seen that the parents – being adults and being in a position where they are capable of making informed decisions where the foetus is not – probably do have good grounds for this difference in treatment. When this is the case however, the issue of universalizability is no longer as relevant, and what is rationally benevolent becomes the more significant question.

⁵⁴⁵ Universalizability seems to only work in favour of strongly suggesting that a difficult pregnancy be continued – this is the only course of action that it would seem ethically acceptable to define as an 'ought'. It seems as though cases in which we might be able to say that a pregnancy *ought* to be ended are extremely rare, and even then this does not seem entirely morally sound. Universalisation would therefore only seem to apply to prenatal decisions in terms of we 'could' universalise either course of action, rather than we 'ought' to – and it could be argued that this is really no kind of universalisation at all. I do argue however that the context of prenatal decision making is a particularly sensitive and difficult one. The universalizability aspect of Rational Benevolence could possibly be made to apply more effectively in other contexts.

Further to this, the question arises whether Rational Benevolence's practical applicability might only be relevant to Western values. That is, we might ask if the absolute/outcome balance within Rational Benevolence still applies in issues of cultural relativism, where certain practices that the Western world would reject on ethical grounds are accepted in certain societies. Again, this is too vast a topic to broach here, but I do suggest that Sidgwick's fundamental Kantian maxim that 'I must be able to judge that what is right for me is right for all other people in this situation, unless there are significantly

Rational Benevolence itself to limit action, it is apparent that the limits imposed here are actually also relatively few – if they can be discerned at all. This is for the reason that was noted in the previous section, that there is a broad spectrum of what can be considered reasonable, or benevolent, or good, in many ethical situations – and it is vital to a liberal society that this range of interpretations is protected. But although Rational Benevolence itself cannot ultimately dictate how good or Happiness should be defined, it is able to specifically and peculiarly account for that wide range of interpretations of the good – and defend them on the basis that they are all *rational*, provided they conform to *some* parameters; the definition of a rationally benevolent action, as given above.

5.4. Conclusions on Rational Benevolence as a Moral Theory

The purpose of the preceding sections has been to assess the capacity of Rational Benevolence as a system of moral guidance – and in terms of Rational Benevolence's value as a theory of normative ethics, it has been seen that major ethical obstacles are encountered (especially in the context of prenatal decision making), when we assess the extent to which Rational Benevolence is capable of identifying individual 'oughts'. Even though there is a case to be made that Rational Benevolence can, in theory, be applied in a normative capacity (extreme decisions that are passed through the criteria of Rational Benevolence do emerge as non-rationally-benevolent), there is still the argument that it ought not to be⁵⁴⁶. In any

different circumstances' could allow for cultural relativism, on the basis of that caveat in the last part of the maxim. This depends, however, on those cultures being viewed in isolation from the norms of other social contexts. When certain practices (such as female genital mutilation (FGM) for example) are viewed from a wider or even a global perspective, it would seem as though the universalising aspect of Rational Benevolence becomes relevant again in the same way as it does for ethically unsound practices such as termination on the basis of eye-colour. *Good* has to be objective to *some* extent, in order to create rationally benevolent actions – and it appears very hard to argue for the *rational* good in a practice such as FGM. Rational Benevolence may, therefore, be able to establish trans-cultural ethical principles.

⁵⁴⁶ The phrase that seems to materialise as the most helpful way to summarise this is that even though what is rational may – according to Rational Benevolence – technically be what is *right*, this might not actually be always what is *best*. This indicates the latent and immovable presence of the 'good' in the human moral experience, and draws attention to a way in which 'right' and 'good' still often represent fundamentally different concepts. I argue however that Rational Benevolence provides an effective means of reconciling 'good' and 'right' to some extent, and in this way accounts precisely for that situation in which 'what is right is not necessarily what is best'. The 'right' in Rational Benevolence is formed of both absolute and non-absolute moral properties, and therefore – as

case, even if this *could* legitimately be done, it is so unlikely that two sets of parents would ever be in exactly the same situation, with exactly the same problems and concerns, that creating an 'ought' is simply an unfeasible and pointless task; it might never apply again, to anyone, anywhere.

But I wish to emphasise here that drawing specific oughts out of Rational Benevolence was not the specific remit of this thesis. My aim was to explain moral ambivalence by arguing that the ambivalence is a natural result of the equal reasonableness of the two types of moral principle we find operating in our moral experiences, and to develop a theory that demonstrates that equal reasonableness, in order that that theory might better justify and support *both* courses of action simultaneously. I argue that Rational Benevolence does offer precisely such a theory. This was clearly demonstrated by its descriptive value. The superficial descriptive value of Rational Benevolence defends the ambivalence by arguing that both the termination or the continuation of a difficult pregnancy *is* (on a *prima facie* basis) morally reasonable, and its substantive descriptive value defends at a deeper level by actively demonstrating how and why each course of action is made morally valid by a necessary interdependence between rational utilitarian and deontological principles, of which Benevolence is comprised. Where its normative value may be somewhat practically limited by the difficulties in establishing universalisable oughts in this context, what is more important about Rational Benevolence is that it universalises *Benevolence* (rather than independent actions) and in so doing makes this Benevolence (or this good, or Happiness) a) rational, and b) widely applicable. By uniting Benevolence with what is rational, the 'good' inherent in the concept is given a far more robust place in this system of the 'right', and in this way, Rational Benevolence actually can effectively protect the vast range of decisions that are being made in very difficult and morally ambivalent circumstances at the practical level. Rationally benevolent actions are both flexible, *and*, to some extent, limited – and this is precisely what is required.

Here, again, Sidgwick is right when he says that the truth lies between two conclusions. In returning to this paradigmatic passage at the end of this discussion on how deontological and utilitarian principles work together in practice, hardly any

argued more extensively above– something would not be 'right' if it did not include reference to both. Ideally, if a course of action did include reference to both, this really would be what is 'best'.

paraphrasing is required: Sidgwick's very words themselves reflect the conclusions I have just drawn. There are certain absolute principles, the truth of which is manifest – but as for what we ought to do in any particular case, particular duties have still to be determined by some other method. This is directly representative of the fine balance between the absolute-principle nature of Benevolence, and its ability to accommodate a range of views within one ultimately rational end.

Now that the mechanisms of Rational Benevolence have been explored from both conceptual and practical angles, it only remains to assess where the theory could possibly apply at an actual societal level, and where and why as a theory it meets certain limitations. These are the concerns of these last two sections, through which the final conclusions on Rational Benevolence will be drawn.

5.4.a. Potential Applications of Rational Benevolence: Prenatal Technologies, Autonomy, and Political Theory

It was seen above that Rational Benevolence meets certain restrictions in its normative capacity, but it was also seen that it does at least indicate where the boundaries on rationally benevolent behaviour might be. It could therefore possibly be used to better support policy that is already in place on issues such as reasonable grounds for termination. i.e. Rational Benevolence may be able to provide a stronger moral-philosophic defence of the current reasons for which termination is allowed⁵⁴⁷. This is the most obviously contentious area of prenatal medicine, and it is here that ethics has most often to strive to justify the decisions being made. As per the above, Rational Benevolence would not be seeking to impose 'oughts' on any grounds for termination. Interestingly however, there has been emerging over the last two decades an area of reproductive ethics which does have the propensity to imply such oughts. This is 'the new eugenics', and its rather more extreme relation, Transhumanism⁵⁴⁸. Here, theorists argue that reproductive

⁵⁴⁷ These reasons are mainly a) When the physical or mental health of the mother is under serious threat, b) When the child of the pregnancy would have an unacceptable quality of life, and c) Social factors, such as extreme poverty, or the parent(s) inability to cope with a(nother) child, leading to a serious risk to the wellbeing of the existing and/or future child.

⁵⁴⁸ The term 'new eugenics' is understood in different ways by different people. For some, it simply refers – without any kind of real ethical subtext – to the use of prenatal medicine to ensure that a child is born as healthy as possible. Of course, non-medical cosmetic selection is also considered to be a form of eugenics, and to this the moral objections are generally stronger. For individuals such as Tom Shakespeare and Ann Kerr however, who write on ethics in this context from the point of view of the

technology should be used for the avoidance of disabilities – with some even arguing that human enhancement is actually a moral duty. This naturally translates into a situation in which termination on the basis of disability becomes morally obligatory. It would seem as though the response of Rational Benevolence to this would be to argue that such an ought could never be imposed; that terminations made on these grounds do not show an adequate balancing of the equal absolute-principle based and outcome based values that make up a rationally benevolent course of action, and that therefore Rational Benevolence must condemn this practice. As was said above, not *all* interpretations of ‘good’ qualify as rationally benevolent. I maintain that arguing against obligatory termination for disability, and against the potential for extreme forms of human enhancement, is the preferable role of Rational Benevolence in this context, and that in this case Rational Benevolence may even be able to impose an ‘ought *not*’ on arguments that claim we should terminate all disability affected pregnancies, thus preventing the proliferation of this practice⁵⁴⁹.

disabled community, the term ‘eugenics’ still carries with it the dangerous connotations of government involvement, and negative societal attitudes (see Shakespeare and Kerr, *Genetic Politics*). According to Shakespeare and Kerr et al, *any* kind of anti-disability prenatal selection is a form of eugenics. This argument is made not just on the basis that more and more people are seeking a ‘good birth’, but together with the more ominous consequences of these choices for society in general; the less people are born with disabilities, the worse the conditions will be for those who are disabled, and the worse the stigma attached. The wider socio-political result of those consequences is a society that sends a message that it does not tolerate disability. This would obviously be a morally unacceptable situation. The concerns of the disabled community in this context are important, and not unfounded. However, the practices to which they object can at least (usually) be defended on medical grounds. The other interpretation of ‘new eugenics’, or ‘liberal eugenics’ as it often called (see Nicholas Agar, *Liberal Eugenics: In Defence of Human Enhancement* (Oxford: Blackwell: 2004), is that selection is done not just for medical purposes, but for specifically *enhancing* purposes. This ‘enhancement’ can include medical aspects such as greater resilience to disease for example, but in its most extreme form enhancement entails Transhumanism – a movement that ultimately envisages a time when humans are born not just free from painful and debilitating conditions, but with the *best* physical and mental attributes possible. Liberal eugenics still aims to minimise state involvement in its theory, given the obvious nefarious results of past governments’ eugenics programs, and Transhumanism is often decried by many as mere science fiction. But the fact of the matter is that the technologies supported by liberal eugenics *are* being advanced, and the possibility of society and government having to deal with the moral management – the use or restriction – of these technologies is actually very real indeed. See John Harris’ provocatively entitled book *Enhancing Evolution: The Ethical Case for Making Better People*, for the best and most concise defence of the Transhumanist position, and the ethics it employs. For other notably Transhumanist positions and discussions, see Nick Bostrom and Julian Savulescu (eds.) *Human Enhancement* (Oxford: Oxford University Press:2009); Ronald Bailey, *Liberation Biology; The Scientific and Moral Case for the Biotech Revolution* (New York: Prometheus Books: 2005); Gregory R. Hansell and William Grassie (eds.) *H+/: Transhumanism and its Critics* (Philadelphia: Metanexus Institute: 2011); Max More and Natasha Vita-More (eds.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future* (Oxford: Wiley-Blackwell: 2013); Ronald Cole-Turner (ed.) *Transhumanism and Transcendence; Christian Hope in an Age of Technological Advancement* (Washington D.C.: Georgetown University Press: 2011).

⁵⁴⁹ I do not mean to assume here that this use of reproductive technology will simply continue to increase unchecked in the absence of Rational Benevolence – there are of course already guidelines

However, Harris has used the language of moral *duty* in relation to our potential to ‘enhance’ the human race. Although Rational Benevolence maintains that Benevolence is the sole duty, as opposed to individual actions, this Benevolence is defined as the general Happiness. Harris’ use of ‘duty’ in this context may well require Rational Benevolence to reconsider whether this general Happiness inherent in Benevolence can in fact be seen to incorporate the apparent ‘benefits’ of enhancement, and if necessary defend itself against this interpretation.

In somewhat less controversial terms of how Rational Benevolence could be societally applied, it is my view that this ability of Rational Benevolence to defend both the termination and the continuation of difficult pregnancies could go a significant way towards strengthening the existing framework of autonomy⁵⁵⁰. It has already been noted in the restatement of the thesis’ main questions above that it is no good simply propounding autonomy itself as the answer to moral ambivalence – autonomy does not always create an understanding space in which parents can make these decisions, and this is, I suggest, largely due to the continual and unrelenting conflict between the two viewpoints. As Sidgwick said, in a passage that is extraordinarily prophetic of many ethical dilemmas that we see in the modern day – ‘if there are different views of the ultimate reasonableness....it is easy to see that any single answer to the question “why” will not be completely satisfactory as it will be given only from one of these points of view, and will always leave room to ask the question from some other’ (I:1:6). If, however, *both* could be argued to be morally reasonable, or, to use the most direct language, if *both* could be argued to be *right*, then clearly both courses of action could be more robustly defended – both theoretically, as was just outlined, and practically, at the level of individual autonomy. This will not, of course, prevent the definitely non-ambivalent from having those views. It would be a huge overestimation of the capacity of Rational Benevolence to claim that it would simply change everyone’s minds, especially on such an ethically

in place that aim to restrict it to within reasonable limits. But it is clear from the volume of literature that there has been a movement towards a liberal eugenics/Transhumanist view of disability, and exactly what is meant by those ‘reasonable limits’ may well have to be re-addressed. It is to this type of discussion that the moral theory of Rational Benevolence could be relevant.

⁵⁵⁰ This is of course an extension of the first point, that Rational Benevolence could be used to better support the existing reasonable grounds for termination. I argue however that focussing on autonomy is probably the more effective method through which this could be done; at the very least, the investigation into Rational Benevolence as a theory of normative ethics would also by extension defend against the introduction of highly questionable grounds for termination, such as those put forwards by extreme Transhumanists.

sensitive topic. But if nothing else, the investigation into Rational Benevolence as a theory of normative ethics demonstrated that individual choice must, generally, be held up as paramount in this context. And if Rational Benevolence was implemented correctly - through policy, through relevant healthcare professionals responsible for working with parents, perhaps even through raising the profile of such a position on social media - a philosophical basis for the equal moral validity of each course of action could provide parents who are faced with such decisions with a helpful line of defence against those who wish to enforce their alternative views⁵⁵¹.

Lastly on this question of Rational Benevolence's societal applicability, I theorise that Rational Benevolence also has the potential to apply to areas outside of prenatal medicine/reproductive technology. Following directly on from the topic of individual autonomy, one such area is political theory – and specifically the contention between individual liberalism and communitarian concerns that arose with Rawls' utilising of social contract theory to defend the notion of individual rights⁵⁵².

⁵⁵¹ I also speculate here that Rational Benevolence could be highly relevant specifically to the area of genetic counselling. It is the task of genetic counsellors to fully inform parents as to the nature of any conditions that are discovered prenatally – as to the severity of the condition, the specific care needs, and the potential outcome. In many cases, genetic counsellors are performing sensitively, and objectively. There is a body of literature however that reports on genetic counsellors who are, whether obviously or covertly, imposing their own views on parents who are already experiencing this highly distressing situation (see this thesis' introduction, pp.9-13) Of course, this will be most pertinent to parents who are ambivalent; parents who are fully decided in continuing the pregnancy still have contact with genetic counsellors in order to prepare them for the future, but any views of those counsellors are unlikely to have an impact. There is perhaps however a case to be made that if parents are only tentatively decided to continue a pregnancy, and they then encounter a genetic counsellor with strong views, this could fundamentally affect their decision. I argue not that Rational Benevolence could simply be presented to healthcare professionals who work in this capacity as some kind of 'holy grail' of how they are to treat their clients, but rather that it could be used at a broader level, to ensure a greater awareness of the ethical value of both courses of action.

⁵⁵² For a comprehensive overview of the individual liberalism vs communitarianism debate of the 1970s, 80s, and 90s, see Ronald Dworkin, 'Liberalism', in, Stuart Hampshire (ed.) *Public and Private Morality* (Cambridge: Cambridge University Press: 1978); Daniel Bell, *Communitarianism and its Critics* (Oxford: Oxford University Press: 1993); Robert B. Thigpen and Lyle A. Downing, 'Liberalism and the Communitarian Critique', in, *American Journal of Political Science*, Vol.31, No.2 (1987); Allen Buchanan, 'Assessing the Communitarian Critique of Liberalism', in, *Ethics*, Vol.99, No.4 (1989), 852-882; Charles Taylor, 'Cross-Purposes; The Liberal-Communitarian Debate', in, Charles Taylor, *Philosophy and the Human Sciences Vol.2* (Cambridge MA: Harvard University Press: 1985); Michael Sandel, *Liberalism and the Limits of Justice* (Cambridge: Cambridge University Press: 1982), and *Liberalism and its Critics* (New York: New York University Press: 1982); Richard Rorty, 'Objectivity, Relativism, and Truth', in, *Philosophical Papers Vol.1* (Cambridge: Cambridge University Press: 1991); Will Kymlicka, 'Liberalism and Communitarianism', in, *Canadian Journal of Philosophy*, Vol.18, No.2 (1988), 181-204; Michael Walzer, *Spheres of Justice: A Defence of Pluralism and Equality* (New York: Basic Books: 1983); Jurgen Habermas 'The Public Sphere', in, Steven Seidman (ed.) *Jurgen Habermas on Society and Politics: A Reader* (London: Beacon Press: 1989); Alasdair MacIntyre, *Whose Justice? Which Rationality?* (London: Gerald Duckworth Press Ltd: 1981), among many others. I wish to point out here that by 'communitarianism', I do not specifically mean the movement of that name created by Amitai Etzioni in the 1980s, which may be referred to as 'responsive communitarianism' and is often criticised for being radically conservative - but rather the position that

The central argument here is that this entrenchment of individual rights is ethically at odds with an interest in the collective good – and the problem is a familiar one to many areas of political philosophy⁵⁵³. It is my view that this contention once again demonstrates the traditional divide between deontological and utilitarian moral principles, with deontological, absolute-principle based concerns, and utilitarian, outcome based concerns representing individual and communitarian interests respectively. I argue further that, like in the moral ambivalence demonstrated by the Parental Predicament, this conflict is caused by the fact that *both* positions have valid and rational grounds –and that *rights* of individuals are being unhelpfully held as being at odds with the *goods* of communities, and vice versa. It would seem therefore as though a social contract should represent neither individual rights nor collective interests exclusively, but an appropriate balance between the two. I theorise that the construction of Rational Benevolence, through which non-consequentialist values are brought into mutual relationship with outcome based values, represents a means which the two types of socio-political concerns could be reconciled.

holds that individuals do not exist in an ‘atomised’ state, but as integral parts of relationships and interactions that are usually informed by the values of the wider social context in which they live. Reference to these relationships and to an individual’s role in the wider community are, it is argued by communitarians, vital to self-understanding. For this reason, emphasis on the importance of private individual autonomy is naturally reduced which is what brings it into conflict with the liberalist ideology that upholds the primacy of the individual, and individual rights.

Walzer, Sandel, Taylor, and MacIntyre especially do not agree with John Rawls’ model for obtaining principles of justice, which Rawls famously argues must be done from behind the veil of ignorance. In this imagined state, individuals have no idea what status or role they themselves will have in that society and the resulting principles therefore reflect a highly egalitarian system founded on individual rights. The communitarian critique of this thought experiment is that it is not possible for individuals to exist in the truly un-related, abstract, ‘atomised’ (Taylor, ‘Cross-Purposes’) state that the veil of ignorance requires, even in principle, and that even if this was possible it would be undesirable on the basis that as soon as we encountered *other* people in society, the principles would not hold.

⁵⁵³ It has perhaps become even more important since the rise of the war on terror, and the problems experienced by governments over how to manage those individuals who pose such threats to communities, but who still do have their own individual rights. Although not current, the case of Abdelbaset Al-Megrahi, the Libyan man found guilty of bombing Pan Am Flight 103 over the Scottish town of Lockerbie in 1988, provides a clear example of individual rights coming into powerful disagreement with a notion of collective justice. Despite being sentenced to life imprisonment, Al-Megrahi was released after serving only eight years of his sentence, on the compassionate grounds that he had terminal cancer. Paradoxically, the political principle of *rights* that had brought about his imprisonment was the very same principle by which he was released. The public protest was severe: Al-Megrahi was considered to be responsible for two hundred and seventy deaths, and yet his own individual welfare had secured his release from a just sentence. Many citizens considered there to have been a flagrant violation of the distress, grief and anger caused by his actions, and of the opinions and moral views that are integral to a wider sense of justice, and of what is right. The political struggle here between individual and collective interests is, I believe, another example of the unhelpful oscillation between the entrenchment of *either* individualism or collectivism exclusively.

Ultimately, the reason for this potential of Rational Benevolence to be applied so broadly is that wherever ethical debate is encountered, it so often reveals the contention to be between absolute-principle based and outcome based methods of ethics. The reason for *this*, I argue, is that morality quite simply *is* comprised of both moral properties. It is this situation that Sidgwick observed, this situation that was revealed by his argument for his own moral principles, and this situation that is embodied by Rational Benevolence. Rational Benevolence is, therefore, in the end intended mainly to apply to moral philosophy – to the discourse that seeks to work on a sort of reconciliatory ‘middle ground’ between deontological and utilitarian theories of morality. But Rational Benevolence, I argue, does not just argue for a ‘middle ground’ that aims to reconcile competing theories by simply allowing them to occupy that space at the same time. As noted in the thesis introduction and in the literature review, W.D. Ross, William Frankena, R.M. Hare, and Derek Parfit have all argued for various alliances between the two types of principle – but none have argued for there being a *necessary* relationship between them, which is made evident at a theoretical/meta-ethical level by Sidgwick, and which is embodied by Rational Benevolence⁵⁵⁴. Rational Benevolence argues that this meeting is actually a type of synthesis which demonstrates both deontological and utilitarian properties to be a part of the whole that is the human moral experience, and reveals that without the values of the other, neither is fully functional. Deontological and utilitarian principles remain *different*, but they are no longer *divided*. For this reason, the potential for Rational Benevolence to apply to any situation in which both deontological and utilitarian values are required appears highly promising.

5.4.b. Limitations to Rational Benevolence

In the thesis introduction, I projected where and why Rational Benevolence would meet certain limits and inadequacies. With the theory now available in full, those projections can be revisited, and assessed in light of this information on the theory’s nature, composition, and functions.

⁵⁵⁴ Derek Parfit’s Triple Theory comes closest to establishing this sort of relationship between deontological and utilitarian principles. But I argue that Rational Benevolence could possibly be made more actively applicable than Triple Theory, in the way I have been describing, and as I outlined in the thesis introduction.

It must firstly be admitted that Rational Benevolence, for all its potential to better defend individual choice, could never truly *solve* moral ambivalence. Directly related to this is the further point that claiming both the ending and the continuing of a difficult pregnancy to be morally reasonable will not necessarily make the actual decision making in this context *easier*. Assuming this would be to grossly overestimate the ability of Rational Benevolence – but more importantly it would also perhaps in itself be unethical. The emotional difficulty of the moral ambivalence involved in decisions such as these can not, and should not, be removed by a simple theory of any kind, even if the parents were to literally refer to such a theory, which is unlikely. If this were possible we would be little more impassive automatons, and in this case would no longer require moral dilemmas to be addressed in the first place. However we might account for it, or explain it, moral ambivalence is an integral part of the human moral experience, precisely for the reasons I have given, which is that very often more than one course of actions appears to be the ‘right’ thing to do – and this will always be distressing. Short of us all becoming Hare’s archangels, ethical debate is always going to continue, and the source of that ethical debate is of course the ambivalence between two apparently reasonable and yet fundamentally conflicting points of view. Whereas I hope that the kind of approach to ethical conflict that I have tried to develop in this thesis might quell the fierce opposition between points of view that often increases the distress for individuals who are already experiencing that very conflict at a private level, I do not presume that ethical dilemmas can or will ever be truly resolved⁵⁵⁵.

Further to this point about the irreconcilability of such conflicts, moral pluralism (or value pluralism) would perhaps argue here that if utilitarian and deontological values are truly equal in the way that they are held to be by Rational Benevolence, then a rational choice between the two courses of action is not actually possible at all⁵⁵⁶. However, although it is important to account for this argument from moral pluralism - relevant as it is to Rational Benevolence’s claim that deontological

⁵⁵⁵Isiah Berlin directly agrees with this. In his account of value pluralism, he is clear that moral values may be equally valid, and yet incompatible, and exist in certain conflict without there appearing to be any resolution (Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford Paperbacks: 1969) p.237). He states that moral conflicts are ‘an intrinsic, irremovable element in human life...these collisions of values are of the essence of what they are, and of what we are’. For the particular context of prenatal decision making, we might want to say that we should never be completely comfortable with the idea of terminating any pregnancy, as justifiable as this course of action often is.

⁵⁵⁶ Rational Benevolence is certainly a pluralist theory if the alternative is ethical monism.

and utilitarian properties of morality are equal in value – several theories have been put forwards that address the incommensurability of values, including the idea that plural values are actually *not* incommensurable in the first place, and it would seem as though Rational Benevolence could offer a contribution of its own to how these values are to be weighed⁵⁵⁷. Nonetheless, there is still the question of exactly *how* Rational Benevolence could be made active in this way. I suggested above that Rational Benevolence could be helpful if implemented as part of current healthcare policy, but there does not appear to be an immediately clear way in which this assimilation could take place.

In terms of the more philosophical challenges to the theory, there is the issue that Rational Benevolence as a normative theory of ethics appears to be somewhat self-defeating. This occurs because on one hand it seems to be able to identify where limits to action might lie, and then on the other hand states precisely on the basis of its own terms that it is not legitimate to apply these limits. It may well be then that Rational Benevolence has superficial descriptive value only, and that its substantive descriptive value cannot really yield a normative system at all. But I maintain that even just this descriptive value – that which Rational Benevolence is saying about the nature of morality – may by itself have real practical worth in better protecting decisions, and making more robust the framework of individual autonomy.

For the rest of the points regarding the possible philosophical challenges faced by Rational Benevolence, we must finally return to Henry Sidgwick, and to the conclusions that have been drawn from his work. There is, I imagine – especially given the vast amount of literature that holds Sidgwick to be a classical utilitarian – one particularly outstanding question, that I have not directly addressed (although I hope my arguments in the thesis have gone some way towards answering it). This

⁵⁵⁷ Thomas Nagel provides a helpful summary of how practical wisdom, for example, could possibly answer the problem of incommensurability: 'Provided one has taken the process of practical justification as far as it will do in the course of arriving at the conflict', he writes, 'one may be able to proceed without further justification, but without irrationality either. What makes this possible is judgement – essentially the faculty Aristotle described as practical wisdom, which reveals itself over time in individual decisions rather than in the enunciation of general principles' (Nagel, *Moral Questions* (Cambridge: Cambridge University Press: 1979) p. 135). Bernard Williams also believes that value comparison is possible, because it is not – as pluralism tends to imply – actually impossible to weigh two rational considerations against each other without first reducing them to 'one common comparison in terms of which they can be compared' (Williams, *Ethics and the Limits of Philosophy* (Cambridge MA: Cambridge University Press: 1985) p.17). It is quite possible, Williams argues, to have competing rationalities – a position with which Rational Benevolence directly agrees. See also James Griffin, 'Incommensurability: What's the Problem?', in, Ruth Chang, *Incommensurability, Incomparability, and Practical Reason* (Cambridge MA: Harvard University Press: 1997).

is: Is Rational Benevolence really just utilitarianism? I admit that the question is a relevant one, and that it might be phrased in a number of pertinent ways. Is saying that 'Benevolence as a class of actions is universalisable' essentially just utilitarianism, for example. Or, is it the case that the Point of View of the Universe simply gives, as Peter Singer would argue, a form of hedonistic utilitarianism, in which ultimately we are to aim at the greatest good for the greatest number? Is the fact that Universal Benevolence is the only self-evident principle truly capable of demanding our moral attention ultimately indicating that all rational moral action is essentially utilitarian in nature? Are all 'rationally benevolent' actions therefore simply utilitarian ones? Is Sidgwick's quest for 'proof' really just evidence that he was aiming to establish the utilitarian principle as superior, a task in which Rational Benevolence is a mere instrument?

In first response to these questions, I maintain that in all cases the answer is no. Sidgwick is clear that his moral theory is comprised of *two* fundamental intuitions – only one of which is utilitarian. The other, he states both in *ME* and in the PD, is Kantian. In Sidgwick's hands that particular intuition retains its quality of non-consequentialist moral rightness – and it is only via this version of rightness, and duty, that Sidgwick is able to establish Universal Benevolence in the first place. Even from the earlier stages of *ME*, it is evident that both deontological and utilitarian approaches to morality are, to Sidgwick, rationally valid. Sidgwick's arguments in I:III for rightness being that which is rational specifically includes both absolute principles and end-based moral positions, and they are posited with very little difference between them. This close affinity, maintained by the shared appeal to Reason, is precisely what is embodied by Rational Benevolence. In that theory, Sidgwick effortlessly and logically combines absolute principles and rational ends to produce a system that is as dependent on the idea of non-consequentialism as it is on a utilitarian idea of Benevolence. And it is *this* theory that provides the *proof* of Sidgwick's utilitarianism. The fact that Sidgwick's utilitarianism required this particular basis as the validation of that utilitarianism – indeed, the fact that it is *derived* from it – is compelling evidence that deontological principles are indispensable. Rational Benevolence is not a mere instrument in the qualification of utilitarianism; it is, in its own right, the essential foundation of the whole moral system that Sidgwick holds up as truth –and for that reason, I argue that it cannot simply be called 'utilitarianism'.

Having said this, as the second response to those questions I venture to ask whether in the end it really matters if Rational Benevolence *is* ultimately found to be a utilitarian theory. Sidgwick does indeed state that his search for clear and ethical intuitions has led him to utilitarianism, but given the importance in that process of Rational Benevolence – and its deontological properties - Sidgwick's confirmation of the utilitarian principle need not detract from that importance. Sidgwick's is such a nuanced interpretation of utilitarianism, one that depends on non-typical-utilitarian interpretations of such fundamental concepts as rightness, that his doctrine emerges as one crucially different to most other versions. If, therefore, it was insisted upon that we call his theory 'utilitarianism', this would be acceptable, provided that it was known as *Sidgwickian Utilitarianism* perhaps, and that this term was recognised as including deontological properties in such a way that they are reconciled to that utilitarianism. This might at least increase the recognition that Sidgwick's account of utilitarianism has something more philosophically significant to offer than a repetition, however sophisticated, of 'the greatest good for the greatest number'.

What would be less satisfactory about 'Sidgwickian Utilitarianism' however is that it allows for the continuation of the term 'utilitarianism'. After so many decades of controversy, debate, and varying popularity, the word would likely retain all those trappings and problems that have come to characterise the doctrine – and would, subsequently, probably still invoke the old rivalry with deontological viewpoints. The crucial details of Sidgwick's particular reconciliatory version would be lost, and the divide between utilitarian and deontological principles of morality would be further perpetuated. It has been one of my most specific aims to argue against this conflict, representing as it does, in my opinion, an inaccurate picture of human morality which Sidgwick's theory does truly serve to eradicate. So where 'Sidgwickian utilitarianism' may be acceptable to some extent, I argue that it is 'Rational Benevolence', and the mutually informing synthesis between deontological and utilitarian moral values therein, that provides a more suitable, more exact, and ultimately more powerful representation of the human moral experience.

5.4.c. Final Comments

When it comes to how ultimately to understand Sidgwick, and his theory of Rational Benevolence, I believe that the truth really does lie between two conclusions.

Sidgwick is a utilitarian, but he is also a deontologist. Rational Benevolence is not standard utilitarianism, because it depends on a foundation of non-consequentialism – but it is not wholly deontological, because what that foundation yields is a utilitarian principle. Deontological moral properties are vital, but insufficient by themselves; utilitarian properties are vital, but unsubstantiated by themselves. Rational Benevolence is objective, but allows for appropriately flexible judgements. Rational Benevolence can recognise both the unconditional aspects of human life, and those places in which life is affected by interpretations of the good. Whichever one of these pairs of statements is chosen, Sidgwick occupies the space between them. That space, identified by myself at the start of this thesis and by Sidgwick at the start of *ME*, is the one in which we appear to agree with a range of very valid but very different assertions about what is morally reasonable. And in developing a theory that accounts for that fact, Sidgwick has created some unity – a position from which we need not adhere to either one view or the other, and a position in which the traditional divide between deontological and utilitarian values dissolves under an acceptance of the rationality of both positions.

I stated in the thesis introduction that Sidgwick hoped to aid towards the construction of a moral system that could bring together the divergent but equally defensible moral beliefs that make up the human moral experience. It is my view that Rational Benevolence does this. Sidgwick himself was not explicit about that fact in *ME* – but I believe that it is not a great leap to suggest that, by the end of his life, Sidgwick did wish to emphasise the affinity between deontological and utilitarian moral principles that he had developed in *ME*. This is, after all, the striking conclusion that he draws twice in the personal document – one of the last documents that Sidgwick was concerned to disclose at the end of his life. Sidgwick's brother in law, Arthur Balfour, always maintained that he believed Henry Sidgwick to have had 'something important left to say', that he hadn't had time to express before his untimely death. In my opinion, the personal document, along with *ME* itself, suggests that this argument for the mutual relationship between deontological and utilitarian principles might have been it. It is my own hope therefore that the work done in this thesis has gone some way to arguing what Sidgwick might have argued himself, had he lived to do so.

BIBLIOGRAPHY

Agar, Nick, *Liberal Eugenics: In Defence of Human Enhancement* (Oxford MA: Blackwell, 2004)

Albee, E., *A History of English Utilitarianism* (Abingdon-on-Thames: Routledge: 1901)

Anscombe, G.E.M., 'Modern Moral Philosophy', in, *Philosophy*, Vol.33, No.124 (1958) 1-19

Audi, Robert, *The Good in the Right* (Princeton: Princeton University Press: 2004)

Bailey, Ronald, *Liberation Biology; The Scientific and Moral Case for the Biotech Revolution* (New York: Prometheus Books: 2005)

Bain, A., 'Mr Sidgwick's Methods of Ethics', in, *Mind*, Vol.1, No.2 (1876) 179-197

Barker, Henry, 'A Recent Criticism of Henry Sidgwick's Methods of Ethics', in *Philosophical Review*, Vol.11, No.6 (1902) 617-603

Barry, B. 'John Rawls and the Search for Stability', in, *Ethics*, Vol.105 (1995) 874-915

Bayles, Michael D. (ed.) *Contemporary Utilitarianism* (New York: Doubleday: 1968)

Baldwin, Thomas *G.E. Moore; The Arguments of the Philosophers* (Oxford: Routledge:1992)

Beanblossom, Ronald, 'Russell's Indebtedness to Reid', in, *The Monist*, Vol.61, No.2 (1978) 192-204

Bell, Daniel, *Communitarianism and its Critics* (Oxford: Oxford University Press: 1993)

Bentham, Jeremy, *Memoirs* (in, John Bowring (ed.) *The Works of Jeremy Bentham Vol.10 [1839]* (London: W.Tait:1843)

Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation* (1789) Ch.1, §4 (in, John Bowring (ed.) *The Works of Jeremy Bentham Vol.1 [1839]* (London: W.Tait:1843)

Berlin, Isaiah, *Four Essays on Liberty* (Oxford: Oxford Paperbacks: 1969)

Blackwell, Kenneth, Brink, Andrew, Griffin, Nicholas (eds.) *The Collected Papers of Bertrand Russell Vol.1* (London: Routledge: 1983) pp.209-10

Blanshard, Brand, *Reason and Goodness* (London: George, Allen, and Unwin: 1961)

Bosanquet, Bernard, 'Review of *Principia Ethica*', *Mind* (Vol.13, 254-261) in 1904

Bostrom, Nick, and Savulescu, Julian, (eds.) *Human Enhancement* (Oxford: Oxford University Press:2009)

Bradley, F.H. 'Mr Sidgwick on Ethical Studies', in, *Mind*, Vol.2, No.122 (1877)

Brandt, Richard, *A Theory of the Good and the Right* (Oxford: Clarendon Press: 1979)

Brandt, Richard, *Morality, Utilitarianism and Rights* (Cambridge: Cambridge University Press: 1992)

Brandt, Richard, *Ethical Theory: The Problems of Normative and Critical Ethics* (Englewood Cliffs, N.J: Prentice-Hall: 1959)

Brandt, Richard, 'A Theory of the Good and the Right'; in, *Philosophical Studies*, Vol.42 (1982) 119-139

Brink, David O., 'Sidgwick and the Rationale for Rational Egoism', , in, Bart Schultz, *Essays on Henry Sidgwick* (Cambridge: Cambridge University Press: 1992) pp.199-240

Brink, David O. ('Common Sense and First Principles in Sidgwick's Methods' in, *Social Philosophy and Policy*, Vol.11, No.1 (1994) 179-201)

Broad, C.D. *Five Types of Ethical Theory* (Paterson, N.J: Littlefield Adams, 1930)

Broad, C.D. 'Critical Notice of W.D. Ross, *Foundations of Ethics*', in, *Mind*, Vol.49 (1940) 228-239

Bryant, John, and Searle, John, *Life in Our Hands: A Christian Perspective on Genetics and Cloning* (Leicester: Inter-Varsity Press: 2004)

Bryce, James, 'Henry Sidgwick', in, *Studies in Contemporary Biography* (London: Macmillan: 1903)

Buchanan, Allen, 'Assessing the Communitarian Critique of Liberalism', in, *Ethics*, Vol.99. No.4 (1989), 852-882

Calderwood, H. 'Mr Sidgwick on Intuitionism', in, *Mind*, Vol.1, No.2 (1876) 197-206

Candlish, Stewart, 'The Truth About F.H. Bradley', in, *Mind*, Vol.98, No.391 (1989) 331-348

Card, Claudia, 'Responsibility Ethics, Shared Understandings, and Moral Communities', in, *Hypatia*, Vol.17, No.1 (2002) 41-155

Carson, Tom, 'Hare on Utilitarianism and Intuitive Morality, in, *Erkenntnis*, Vol.39, No.3 (1993) 305-331

Cocking, Dean, and Oakley, Justin, 'Indirect Consequentialism, Friendship, and the Problem of Alienation, in, *Ethics*, Vol.106, No.1 (1995) 86-111

Cole-Turner, Ronald (ed.) *Transhumanism and Transcendence; Christian Hope in an Age of Technological Advancement* (Washington D.C.: Georgetown University Press: 2011)

Collini, Stefan, 'Idealism and "Cambridge Idealism"', in, *The Historical Journal*, Vol.18, No.1 (2009) 171-177

Collini, Stefan, ('My Roles and their Duties: Sidgwick as a Philosopher, Professor, and Public Moralist', in *Proceedings of the British Academy*, Vol. 109 (2001), 9-49

Crisp, Roger, *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics* (Oxford: Oxford University Press: 2015)

Crisp, Roger, 'Sidgwick and the Boundaries of Intuitionism', in, Stratton-Lake, Phillip (ed.) *Ethical Intuitionism* (Oxford: Clarendon Press: 2002) 56-75

Crisp, Roger, *Underivative Duty: British Moral Philosophers from Sidgwick and Ewing* (Oxford: Oxford University Press: 2011)

Crisp, Roger, *British Ethical Theorists from Sidgwick to Ewing* (Oxford: Oxford University Press: 2014)

D'Amico, Robert, 'R.M. Hare, 1919-2002', in, *Proceedings and Addresses of the American Philosophical Association*, Vol.76, No.2 (2002) 129-130

Daurio, Janice, 'Sidgwick on Moral Theories and Common Sense Morality, in, *History of Philosophy Quarterly*, Vol.14, No.4, (1997) 425-445

Dawes, Ann, *Henry Sidgwick* (Cambridge: Biograph: 2007)

De Lazari-Radek, Katarzyna, 'Sidgwickian Ethics' (Review), *Mind*, Vol.123, No. 491 (2014), 951-956

Dewey, C.J. 'Cambridge Idealism'; Utilitarian Revisionists in Late Nineteenth-Century Cambridge', in, *The Historical Journal*, Vol.17, No.1 (1974) 63-78

Donagan, Alan, 'Sidgwick and Whelwellian Intuitionism; Some Enigmas', in, *Canadian Journal of Philosophy*, Vol.7, No.3 (1977) 447-465

Dworkin, Ronald, 'Liberalism', in, Stuart Hampshire (ed.) *Public and Private Morality* (Cambridge: Cambridge University Press: 1978)

Ekberg, Merryn, 'Maximising the Benefits and Minimising the Risks Associated with Prenatal Genetic Testing', in, *Health, Risk, and Society*, Vol.9, No.1 (2007), 67-81

Ekberg, Merryn, 'Second Opinions: The Old Eugenics and the New Genetics Compared', in, *Social History of Medicine*, Vol.20, No.3 (2007) 581-593

Ewing, A.C. *Second Thoughts in Moral Philosophy* (New York: Macmillan:1959)

Fitzpatrick, Tony, 'Before the Cradle: New Genetics, Biopolicy, and Regulated Eugenics', in, *Journal of Social Policy*, Vol.30, No.4 (2001) 589-612

Foot, Philippa, *Virtues and Vices* (Oxford: Blackwell: 1978)

Foot, Philippa, 'The Problem of Abortion and the Doctrine of Double Effect', in, *Virtues and Vices and other Essays in Moral Philosophy Vol.19* (Oxford: Blackwell Publishers: 1978)

Frankena, William, *Ethics* (Englewood Cliffs, N.J.: Prentice-Hall, Inc.:1963)

Frankena, William, 'Sidgwick and the Dualism of Practical Reason', in *The Monist*, Vol.58 (1974) 449-467

Frey, R.G. 'Act Utilitarianism: Sidgwick, Bentham, or Smart?', in, *Mind*, Vol.86, No.341 (1977)

Galton, David, *Eugenics; The Future of Human Life in the 21st Century* (London: Abacus, 2001)

Garner, Richard T. and Rosen, Bernard, *Moral Philosophy: A Systematic Introduction to Normative Ethics and Meta-ethics* (New York: Macmillan:1967), p.70)

Gaus, Gerald F. and Kukathas, Chandran, (eds.) *Handbook of Political Theory* (London: SAGE Publishing:2004) pp.413-417

Glover, Jonathon, *What Sort of People Should There Be?* (London: Penguin Books, 1984)

Glover, Jonathon, *Choosing Children; Genes, Disability, and Design* (Oxford: Clarendon Press, 2006)

Goodin, Robert E. *Utilitarianism as a Public Philosophy* (Cambridge University Press:1995)

Gray, John, 'Indirect Utility and Fundamental Rights', in, Gray, *Liberalisms: Essays in Political Philosophy* (London: Routledge: 1989)

Griffin, James, 'Incommensurability: What's the Problem?', in, Chang, Ruth, *Incommensurability, Incomparability, and Practical Reason* (Cambridge MA: Harvard University Press: 1997)

Griffiths, A. Phillips, 'Review of *Generalisation in Ethics*', in, *Philosophical Books*, Vol.3, No.1 (1962) 18-21

Habermas, Jurgen, *Truth and Justification* (Cambridge MA: MIT Press: 2003)

Habermas, Jurgen, 'The Public Sphere', in, Steven Seidman (ed.) *Jurgen Habermas on Society and Politics: A Reader* (London: Beacon Press: 1989)

Hager, Paul, *Continuity and Change in the Development of Russell's Philosophy* (London: Kluwer Academic Publishers: 1994)

Hall, Everett, 'The Philosophy of G.E. Moore', in, *Philosophical Review*, Vol. 53, No.1 (1944) 62-68

Hansell, Gregory R. and Grassie, William (eds.) *H+/: Transhumanism and its Critics* (Philadelphia: Metanexus Institute: 2011)

Hare, R.M. *Moral Thinking* (Oxford: Clarendon Press: 1981)

Hare, R.M. *Sorting out Ethics* (Oxford: Clarendon Press: 2000)

Hare, R.M. 'Rules of War and Moral Reasoning', in, *Philosophy and Public Affairs*, Vol.1 No. 2 (1972) pp. 166 -181

Harris, John, 'Clones, Genes, and Immortality', in, *Wonderwoman and Superman* (Oxford: Oxford University Press: 1998)

Harris, John, *Enhancing Evolution; The Ethical Case for Making Better People* (Princeton: Princeton University Press: 2007)

Harris, John 'Enhancements are a Moral Obligation', in, Bostrom, Nick, and Savulescu, Julian, (eds.), *Human Enhancement* (Oxford: Oxford University Press, 2011)

Harrison, Ross, 'Cambridge Philosophers IV: Henry Sidgwick', in, *Philosophy*, Vo.71, No.277 (1996) 423-438

Harrison, Ross (ed.), *Henry Sidgwick* (Oxford: Oxford University Press: 2001)

Harrison, Ross 'The Sanctions of Utilitarianism', in *Proceedings of the British Academy*, Vol. 109 (2001), 93-116

Harsanyi, John C. 'Rule Utilitarianism and Decision Theory', in, *Erkenntnis*, Vol.11, No.1 (1977) 25-53

Harsanyi, John C. 'Morality and the Theory of Rational Behaviour', in, Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond*, (Cambridge: Cambridge University Press) 56-61

Hayward, F.H. 'Mr Hayward's Evaluation of Professor Sidgwick's Ethics: A Reply', in, *Mind*, Vol.11, No.3 (1901) 360-365

Hayward, F.H. 'The True Significance of Sidgwick's Ethics', in *International Journal of Ethics*, Vol.11, No.2 (1901) 175-187

Hills, Alison, 'The Significance of the Dualism in Practical Reason', in, *Utilitas*, Vol.15, No.3 (2003) 315-329

Holley, David, 'Sidgwick's Problem', in, *Ethical Theory and Moral Practice*, Vol. 5, No.1 (2002), 5-65

Hooker, Brad, *Ideal Code, Real World: A Rule Consequentialist Theory of Morality* (Oxford: Clarendon Press: 2002)

Hooker, Brad, 'Rule-Consequentialism, Incoherence, Fairness', in, *Proceedings of the Aristotelian Society*, New Series, Vol.95 (1995) 19-35

Hooker, Brad, 'Ross-style Pluralism versus Rule-Consequentialism', in, *Mind*, Vol.105, No.420 (1996), 531-552

Hooker, Brad, 'Rule Consequentialism', in, *Mind*, New Series, Vol.99, No.393 (1990) 67-77

Hooker, 'Sidgwick and Common Sense Morality', in, *Utilitas*, Vol.12, No.3 (2000) 347-360

Hooker, Brad, 'Must Kantian Contractualism and Rule-Consequentialism Converge?', in, *Oxford Studies in Normative Ethics*, Vol.4 (2014) 34-52

Hooker, Brad, 'Rule Consequentialism', The Stanford Encyclopedia of Philosophy (Winter 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2015/entries/consequentialism-rule/>.

Hooker, Brad, Mason, Elinor, and Miller, Dale E. (eds.) *Reflective Equilibrium and Rule Consequentialism*, in *Morality, Rules and Consequences* (Edinburgh: Edinburgh University Press Ltd: 2000)

Huemer, Michael ('Singer's Unstable Meta-ethics', in, *Singer Under Fire*, pp. 359-379)

Hudson, W.D, *Modern Moral Philosophy* (London: Macmillan:1970)

Hurka, Thomas, 'Sidgwick on Consequentialism and Deontology', in, *Utilitas*, Vol. 26, No.2 (2014) 129-152

Hurka, Thomas, 'Moore in the Middle', in, *Ethics*, Vol.113, 599-628 (2003)

Hurka, Thomas, *British Ethical Theorists from Sidgwick to Ewing* (Oxford: Oxford University Press: 2014)

Hutchinson, Roma, Index to *The Collected Stories of Bertrand Russell*. Russell: The Journal of Bertrand Russell Studies, [S.I.], v. 22, jun. 2002. ISSN 1913-8032. Available at: <https://escarpmentpress.org/russelljournal/article/view/2017>

Irvine, A.D., 'Russell on Method', in, Link, Godehard (ed.) *One Hundred Years of Russell's Paradox* (Berlin and New York: Walter de Gruyter: 2004)

Irvine, A.D. 'Epistemic Logicism and Russell's Regressive Method', in, *Philosophical Studies*, Vol.55 (1989) 303-327

Jacobson, D. 'J.S.Mill and the Diversity of Utilitarianism', in, *Philosopher's Imprint*, Vol.3, No.2 (2003)

James, D.G. *Henry Sidgwick: Science and Faith in Victorian England*, (Oxford: Oxford University Press: 1970)

Jamieson, Dale (ed.) *Singer and his Critics* (Oxford: Blackwell Publishers: 1999)

Jackson, Frank, 'Non-cognitivism, Normativity, Belief', in, *Ratio*, Vol.12, No.4, (1999) 20-435

Jarvis Thompson, Judith, 'The Trolley Problem', in, *The Yale Law Journal*, Vol. 94, No.6 (1984) 1395-1415

Jones, E.E.C. 'Mr Moore on Hedonism', in *International Journal of Ethics*, Vol.16, No.4 (1906) 429-464

Joseph, H.W.B. *Some Moral Problems* (Oxford: Oxford University Press: 1931)

Kagan, Shelley, *The Limits of Morality* (Oxford: Clarendon Press: 1989)

Kant, Immanuel, *Critique of Pure Reason*, 1781

Kant, Immanuel, *Groundwork for the Metaphysics of Morals*, 1785

Kant, Immanuel, *Critique of Practical Reason*, 1788

Kass, Leon, 'The Wisdom of Repugnance' (in, *The New Republic*, Vol.216, No.2 (1997) 17-26).

Keyt, David, 'Singer's Generalisation Argument', in, *Philosophical Review* Vol.72 (1963) 466-476

Kevles, Daniel, *In the Name of Eugenics; Genetics and the Uses of Human Heredity* (Cambridge, MA: Harvard University Press, 1995)

Kirkeby, Eivind, 'Against Moral Intuitions',
([https://www.duo.uio.no/bitstream/handle/10852/24997/Against_Moral_Intuitions.pdf](https://www.duo.uio.no/bitstream/handle/10852/24997/Against_Moral_Intuitions.pdf?sequence=1)
?sequence=1 (2009) accessed 8th July 2016)

Kitcher, Phillip, *The Lives to Come: The Genetic Revolution and Human Possibilities* (New York: Simon and Schuster, 2001)

Kraut, Richard, *What is Good and Why* (Cambridge, MA: Harvard University Press: 2007)

Kuhse, Helga (ed.), *Unsanctifying Human Life* (Oxford: Blackwell Publishers Ltd: 2002)

Kymlicka, Will, 'Liberalism and Communitarianism', in, *Canadian Journal of Philosophy*, Vol.18, No.2 (1988), 181-204

Lacey, A.R. in 'Sidgwick's Ethical Maxims '(*Philosophy*, Vol.34, No. 130 (1959), 217-228

Langenfus, William, ('The Impossibility of Sidgwick's Proof', in, *History of Philosophy Quarterly*, Vol.17, No.1 (2000), 99-120

Levine, S. 'Truth and Moral Validity: On Habermas' Domesticated Pragmatism', in, *Constellations*, Vol.18, No.2 (2011) 244-259

Lyons, David, *Forms and Limits of Utilitarianism* (Oxford: Oxford University Press:1965)

McDonald, Michael, 'Critical Notice', in, *Canadian Journal of Philosophy*, Vol.12, No.2 (1982) 375-389

MacIntyre, Alasdair, '*Having Survived Academic Moral Philosophy of the Twentieth Century*', in, O'Rourke, Fran (ed.) *What Happened in and to Moral Philosophy in the Twentieth Century* (Notre Dame, Indiana: Undpress: 2013)

MacIntyre, Alasdair, *A Short History of Ethics* (Abingdon-on-Thames: Routledge & Kegan Paul Ltd: 1967)

MacIntyre, Alasdair, *Whose Justice? Which Rationality?* (London: Gerald Duckworth Press Ltd: 1981)

MacIntyre, Alasdair, *After Virtue* (London: Duckworth: 1985)

Mackie J.L., 'Sidgwick's Pessimism', in, *Philosophical Quarterly*, Vol.26, No.105 (1976) 317-27

McMahan, Jeff, 'Moral Intuition', in, LaFollete, Hugh (ed.) *The Blackwell Guide to Ethical Theory* (Oxford: Blackwell Publishers, 2000) pp.92-110

Margolis, Joseph, 'G.E. Moore and Intuitionism', in, *Ethics*, Vol.87, No.1 (1976) 35-48

Mill, J.S., *Utilitarianism* (Oxford: Blackwell Publishing Ltd: 2003)

Miller, Dale E. 'Hooker's Use and Abuse of Reflective Equilibrium', in Hooker, Mason, and Miller,(eds.) pp.156-178

Milo, Ronald D., 'Bentham's Principle' (*Ethics*, Vol.84, No.2 (1979) 128-139

Moore, G.E., *Principia Ethica* (Cambridge: University Press: 1903)

More, Max, and Vita-More, Natasha (eds.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future* (Oxford: Wiley-Blackwell: 2013)

Moody-Adams, Michele, 'Review of Schultz (ed.) *Essays on Henry Sidgwick*, in, *Victorian Studies*, Vol.37, No.1 (1993) 149-150

Montague, William Pepperell, *The Ways of Knowing, or The Methods of Philosophy* (London: Compton Printing Works Ltd: 1925)

Mulgan, Tim, *The Demands of Consequentialism*, (Oxford: Oxford University Press: 2005)

Murphy, Arthur, *The Theory of Practical Reason* (LaSalle, Ill: Open Court: 1964)

Nagel, Thomas, *Moral Questions* (Cambridge: Cambridge University Press: 1979)

Nakano-Okuno, Mariko, 'Sidgwick and Kant': On the So-Called Discrepancies Between Utilitarian and Kantian Ethics', in Buculo, P., Crisp, R., and Schultz, B. (eds.) *Henry Sidgwick: Happiness and Religion* (Catania: Catania University Press: 2007) 260-332

Nakano-Okuno, Mariko, *Sidgwick and Contemporary Utilitarianism* (London: Palgrave Macmillan: 2011)

Parfit, Derek, *Reasons and Persons* (New York: Oxford University Press: 1984)

Parfit, Derek, *On What Matters* (New York: Oxford University Press: 2011).

Parker, M.H., Forbes, K.L. and Finlay, I. 'Eugenics or Empowered Choice; Community Issues Arising from Prenatal Testing', in, *Australian and N. Zealand Journal of Obstetrics and Gynaecology*, Vol. 42, No.1 (2003) 10-4

Pettit, Phillipm and Kukathas, Chandran, *Rawls: 'A Theory of Justice' and its Critics* (Cambridge: Polity Press: 1980)

Phillips, David, 'Sidgwick, Dualism, and Indeterminacy in Practical Reason', in, *History of Philosophy Quarterly*, Vol.15, No.1 (1998), 57-78).

Phillips, David, *Sidgwickian Ethics* (Oxford: Oxford University Press: 2011)

Pojman, Louis, *Ethical Theory: Classical and Contemporary Readings* (Belmont, CA: Wadsworth:1994)

Pollock, Lansing, 'A Dilemma for Singer', in, *Philosophical Studies*, Vol.33, No.4 (1978), 425

Popper, Karl, *The Open Society and its Enemies* (Oxford: Routledge and Kegan Paul Ltd: 1966)

Quinn, Warren, 'Action, Intentions, and Consequences; The Doctrine of Doing and Allowing', in, *Philosophical Review* Vol.98, No.3 (1989) 355-382

Quinton, Antony, *Utilitarian Ethics* (London: The Macmillan Press Ltd.: 1973)

Rabino, Isaac, 'Genetic Testing and its Implications: Human Genetics Researchers Grapple with Ethical Issues', in, *Science, Technology and Human Values*, Vol. 28, No.3 (2003) 365-402

Raphael, D.D., 'Sidgwick on Intuitionism', in, *The Monist*, Vol.58, No.3 (1974) 405-419

Rashdall, Hastings, 'Professor Sidgwick's Utilitarianism', in, *Mind* (1885) Vol.10, No.38 (1885) 200-226

Rashdall, Hastings, ('Professor Sidgwick's Utilitarianism', in, *Mind*, Vol.10, No.38 (1885) 200-226

Rashdall, Hastings, *The Theory of Good and Evil* (Oxford: Clarendon Press),

Rawls, John, Preface to Sidgwick's *The Methods of Ethics*, 7th Edition (Indianapolis: Hackett Publishing: 1980)

Rawls, John, 'Two Concepts of Rules', in, *The Philosophical Review* Vol.64 (1955) 3-32

Rawls, John, (Herman, Barbara (ed.) *Lectures on the History of Moral Philosophy* (Cambridge, MA: Harvard University Press: 2000)

Rawls, John, *Political Liberalism* (New York: Colombia University Press: 1993

Rawls, John, *A Theory of Justice* (Cambridge MA: Belknap Press of Harvard University Press:1971)

Roberts, Christy, Stough, Laura, and Parrish, Linda, 'The Role of Genetic Counselling in the Elective Termination of Pregnancies Involving Disabilities', in, *The Journal of Special Education*, Vol.36, No.1 (2002) 48-55

Ross, W.D. *The Right and the Good* (Oxford: Oxford University Press: 1930)

Ross, W.D. *The Foundations of Ethics* (Oxford: Oxford University Press: 1939)

Rorty, Richard, 'Objectivity, Relativism, and Truth', in, *Philosophical Papers Vol.1* (Cambridge: Cambridge University Press: 1991)

Russell, Bertrand, *My Philosophical Development* (London: Allen and Unwin: 1959)

Russell, Bertrand, 'On the Foundations of Ethics', in, Pigden, Charles, (ed.), *Russell on Ethics* (London: Routledge: 1999), 206-11

Russell, Bertrand, 'Knowledge by Acquaintance and Knowledge by Description', in *Proceedings of the Aristotelian Society*, Vol.11 (1911) 108-128

Russell, Bertrand, *The Problems of Philosophy* (Oxford: Oxford University Press: 1912)

Russell, Bertrand, *Human Knowledge: Its Scope and Limits* (London: George Allen and Unwin: 1948)

Sandel, Michael, *Liberalism and the Limits of Justice* (Cambridge: Cambridge University Press: 1982)

Sandel, Michael, *Liberalism and its Critics* (New York: New York University Press: 1982)

Sandel, Michael, *The Case Against Perfection: Ethics in the Age of Genetic Engineering* (Cambridge MA: Harvard University Press: 2007)

Santalahti, Paivi, Hemminki, Elini, Latikka, Anne-Marie, and Ryyananen, Marrku, 'Women's Decision-making in Prenatal Screening', in *Social Science and Medicine*, Vol. 46, No.8 (1998) 1067-1076

Savulescu, Julian, 'Why We Should Select The Best Children', in, *Bioethics*, Vol. 15, No.5/6 (2001) 413-426

Scarre, Geoffrey, *Utilitarianism* (London: Routledge: 1996)

Schaler, J.A. (ed.) *Peter Singer Under Fire* (Chicago and La Salle, Ill: Open Court: 2009)

Scheffler, Samuel, 'Morality and Reasonable Partiality', in, Feltham, Brian, and Cottingham, John (eds.) *Partiality and Impartiality: Morality, Special Relationships, and the Wider World* (Oxford: Oxford University Press: 2010)

Schneewind, Jerome B. *Sidgwick's Ethics and Victorian Moral Philosophy* (Oxford: Oxford University Press: 1977)

Schneewind, J.B. 'Sidgwick and the Cambridge Moralists', in, *The Monist*, Vol. 58, No.3 (1974) 371-404

Schneewind, J.B. 'First Principles and Common Sense Morality in Sidgwick's Ethics' (in, *Archiv Für Geschichte der Philosophie* Vol.45, No.2 (1963) 137-156

Schneewind, J.B. 'Korsgaard and the Unconditional in Morality', in, *Ethics* Vol.109, No.1 (1998) 36-48

Schultz, Bart, *Henry Sidgwick: Eye of the Universe*, (Cambridge: Cambridge University Press: 2012)

Schultz, Bart, 'Review of Singer and Lazari-Radek, *Point of View of the Universe: Sidgwick and Contemporary Ethics*', University of Notre Dame, 7/7/2014).

Schultz, Bart, *A more Reasonable Ghost*, <https://roundedglobe.com/books/34a3e7ff-778f-48d5-bca0-ed4e10132715/A%20More%20Reasonable%20Ghost:%20Further%20Reflections%20on%20Henry%20Sidgwick%20and%20the%20Irrationality%20of%20the%20Universe/>

Schlipp, P.A. (ed.) *The Philosophy of Bertrand Russell* (The Library of Living Philosophers Vol.5) (Evanston and Chicago: Northwestern University: 1944).

Seavilleklein, V. 'Challenging the Rhetoric of Choice in Prenatal Screening', in, *Bioethics*, Vol. 23, No.1 (2009) 68-77

Seth, James, 'The Ethical System of Henry Sidgwick' in *Mind*, Vol.10, No.38 (1901) 172-187

Seth, James, 'Reviewed Work: Lectures on the Ethics of T.H. Green, Mr Herbert Spencer and J. Martineau', in, *The Philosophical Review*, Vol.12, No.5 (1903) 548-553

Shakespeare, Tom, and Kerr, Anne (eds.) *Genetic Politics: From Eugenics to Genome* (Cheltenham: New Clarion Press, 2002)

Shakespeare, Tom, 'Losing the Plot'? Medical and Activist Discourses of Contemporary Genetics and Disability, in, *Sociology of Health and Illness*, Vol.21, No.5 (1999) 669-688

Shapria, Amos, 'Wrongful Life' Lawsuits for Faulty Genetic Counselling', in, *Journal of Medical Ethics*, Vol.24, No.6 (1998) 369-375

Shaver, Robert, 'Sidgwick's Minimal Metaethics', in, *Utilitas*, Vol.12, No.3 (2000) 261

Shaver, Robert, 'Sidgwick on Moral Motivation', in, *Philosopher's Imprint*, Vol.6, No.1 (2006) 1-14

Shaver, Robert, *Rational Egoism: A Selective and Critical History* (Cambridge: Cambridge University Press: 1999)

Shaver, Robert, 'Sidgwick's False Friends' (in, *Ethics* Vol.107, No.2 (1997) 314-320).

Shaw, William, *Contemporary Ethics: Taking Account of Utilitarianism* (Oxford: Blackwell: 1999)

Sidgwick, Henry, *The Methods of Ethics* (London: Macmillan: 1874)

Sidgwick, Henry, *Principles of Political Economy* (London: Macmillan and Co.) in 1883

Sidgwick, Henry, *The Elements of Politics* (London: Macmillan and Co.) in 1891

Sidgwick, Henry, *Practical Ethics; A Collection of Addresses and Essays* (London: S. Sonnonschein) in 1898

Sidgwick, Henry, 'Professor Calderwood On Intuitionism in Morals' (*Mind*, Vol.1, No.4 (1876) 563-566)

Sidgwick, Henry, *Outlines of the History of Ethics* (London: Macmillan: 1886)

Sidgwick, Henry, 'Some Fundamental Ethical Controversies' (*Mind*, Vol.14, No.56 (1889) 473-87)

Sidgwick, Arthur, and Sidgwick, Eleanor Mildred, *Henry Sidgwick: A Memoir* (London: Macmillan:1906)

Sidgwick, Henry, 'Green's Ethics', in *Mind*, Vol.9, No.34 (1884) 169-187

Sidgwick, Henry, *Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays*, (London: Macmillan: 1905) 220-221

Sidgwick, Henry 'The Establishment of Ethical First Principles', in, *Mind* Vol.4 (1879), 106-11

Sidgwick, Henry, 'Pleasure and Desire', in, *The Contemporary Review*, Vol. 19 (1872), 662-72

Sidgwick, Henry, *Review of Green and Grose's Edition of Hume's Treatises* (Academy, 30th May, 1874)

Sidgwick, Henry, (Jones, E.E.C (ed.), *Lectures on the Ethics of T.H.Green, Mr Herbert Spencer and J. Martineau* (London: Macmillan: 1902), pp.11-14, 15-22

Skelton, Antony, 'Utilitarian Practical Ethics: Sidgwick and Singer' (in, Bucolo, Crisp, and Schultz (eds.) *Henry Sidgwick: Ethics, Psychics, and Politics* (Catania: University of Catania Press: 2011)

Skelton, Antony, 'Henry Sidgwick's Practical Ethics: A Defense', in, *Utilitas* Vol.8, No.3 (2006) 199-217

Skelton, Antony, 'Schultz's Sidgwick', in, *Utilitas*, Vol.19, No.1 (2007) 91-103

Skelton, Antony, 'Henry Sidgwick's Moral Epistemology', in, *Journal of the History of Philosophy*, Vol.48, No.4 (2010) 491-519

Skelton, Antony, 'Sidgwick's Philosophical Intuitions', in, *Ethics and Politics*, Vol.10, No.2 (2008) 185-209

Skelton, Antony, 'Henry Sidgwick: 1838-1900' (in, Sweet, William (ed.), *Biographical Encyclopedia of British Idealism* (London: Continuum: 2010)

Skelton, Antony, ('Henry Sidgwick', in, Mander, J. and Sell, A. P. F. (eds.) *The Dictionary of Nineteenth Century British Philosophers* (London: Bloomsbury: 2002).

Skelton, Antony, 'On Henry Sidgwick's 'My Station and its Duties'', in, *Ethics*, Vol.125 (2014) 586-591

Skelton, Antony, « Sidgwick's Argument for Utilitarianism and his Moral Epistemology: A Reply to David Phillips », *Revue d'études Benthamiennes* [En Ligne], 12 | 2013, mis en ligne le 10 Décembre 2013 URL : <http://etudes-benthamiennes.revues.org/675>

Skorupski, John, 'Spencer and the Moral Philosophers', in, Francis, Mark, and W. Taylor, Michael (eds.) *Herbert Spencer; Legacies* (Abingdon-on-Thames: Routledge: 2011)

Skorupski, John, *English Language Philosophy 1750-1945* (Oxford: Oxford University Press:1993)

Skorupski, John, 'Sidgwick and the Many-Sidedness of Ethics', in Bucolo, Crisp and Schultz (eds.) *Henry Sidgwick: Happiness and Religion* (2004), 410-422

Singer, Marcus (ed.), *Henry Sidgwick: Essays on Ethics and Method* (Oxford: Oxford University Press: 2000)

Singer, Marcus, 'Common Sense and Paradox in Sidgwick's Ethics', in, *History of Philosophy Quarterly* Vol.3, No.1 (1986) 65-78

Singer, Marcus, 'Sidgwick and Nineteenth-century British Ethical Thought', in Schultz (ed.) *Essays on Henry Sidgwick*, pp.65-92

Singer, Marcus, 'The Many Methods of Sidgwick's Ethics', in, *The Monist*, Vol.58, No.3 (1974) 420-448

Singer, Marcus, 'The Methods of Justice: Reflections on Rawls', *Journal of Value Enquiry*, Vol. 10, No.4 (1976), 286-316

Singer, Marcus, *Generalisation in Ethics* (New York: Alfred A. Knopf: 1955)

Singer, Peter, *How Are We To Live?* (Oxford: Oxford University Press: 1997)

Singer, Peter, and de Lazari-Radek, Katarzyna, *The Point of View of the Universe: Sidgwick and Contemporary Ethics* (Oxford: Oxford University Press: 2014)

Singer, Peter, *Animal Liberation* (London: Pimlico, Random House: 1975)

Singer, Peter, 'Sidgwick and Reflective Equilibrium', in, *The Monist*, Vol.58, No.3 (1974) 490-517

Singer, Peter, 'Ethics and Intuitions', in, *Journal of Ethics*, Vol.9, No.3-4 (2005) 331-352

Singer, Peter, *The Expanding Circle: Ethics, Evolution, and Moral Progress* (Princeton: Princeton University Press: 2011)

Singer, Peter, *Practical Ethics* (Cambridge: Cambridge University Press: 1993)

Sinnot-Armstrong, Walter, 'R.M. Hare, (1919-)

<http://sites.duke.edu/wsa/papers/files/2011/05/wsa-rmhare2001.pdf> (accessed 18/6/2016)

Smart, J.J.C., 'Extreme and Restricted Utilitarianism', in, *The Philosophical Quarterly*, Vol.6, No.25 (1950) 344-354

Smart, J.J.C. *An Outline of a System of Utilitarian Ethics* (Carlton: Melbourne University Press: 1961)

Smart, J.J.C. and Williams, Bernard, *Utilitarianism; For and Against* (Cambridge: Cambridge University Press:1973)

Snare, Frank, 'John Rawls and the Methods of Ethics', in, *Philosophy and Phenomenological Research* Vol.36, No.1 (1975) 100-112

Sobel, Howard, 'Generalisation Arguments', in, *Theoria* Vol.31, No.1 (1965) 32-60

Sorley, W.R., *A History of English Philosophy* (New York: G.P Putman's Sons: 1921)

Stephens, Lesley, 'Sidgwick's Methods of Ethics' in *Fraser's Magazine* (1875) n.s.11

Stratton-Lake, Phillip, *Ethical Intuitionism: Re-Evaluations*, (Oxford: Oxford University Press: 2002).

Stratton-Lake, Phillip (ed.), *The Right and the Good* (Oxford: Oxford University Press: 2002)

Sturgeon, Nicholas, 'Ethical Intuitionism and Ethical Naturalism', in, Phillip Stratton-Lake (ed.) *Ethical Intuitionism* (Oxford: Clarendon Press: 2002) pp.184-211

Sverdlik, Steven, 'Sidgwick's Methodology', *Journal of the History of Philosophy* Vol. 23, No.4 (1985), 537-553

Taylor, A.E. 'Review: *Philosophy: Its Scope and Relations: Henry Sidgwick*', in, *Ethics*, Vol.13, No.3 (1903) 377

Taylor, Charles, 'Cross-Purposes; The Liberal-Communitarian Debate', in, Taylor, Charles, *Philosophy and the Human Sciences Vol.2* (Cambridge MA: Harvard University Press: 1985)

Thigpen, Robert B. and Downing, Lyle A. 'Liberalism and the Communitarian Critique', in, *American Journal of Political Science*, Vol.31, No.2 (1987) 637-655

Tooley, Michael, 'Personhood', in, Peter Singer and Kuhse (eds.) *Companion to Bioethics* (Oxford: Blackwell Publishing: 2009)

Tropman, Elizabeth, 'Renewing Moral Intuitionism', in, *Journal of Moral Philosophy*, Vol.6, No.4 (2009) 440-463

Walzer, Michael, *Spheres of Justice: A Defence of Pluralism and Equality* (New York: Basic Books: 1983)

Warnock, G.J., *Contemporary Moral Philosophy* (New York: St. Martin's Press: 1967)

Weinstein, D. *Utilitarianism and the New Liberalism* (Cambridge: Cambridge University Press: 2011)

Wiggins, David, *A Sensible Subjectivism* (Oxford: Oxford University Press:1998).

Williams, Bernard, *Ethics and the Limits of Philosophy* (London: Fontana: 1985)

Williams, Bernard, 'The Point of View of the Universe', in, Williams, Bernard, *Making Sense of Humanity and Other Philosophical Papers 1982-1993* (Cambridge: Cambridge University Press: 1995)

Wolfe, Tom, 'The 'Me' Decade and the Third Great Awakening', in *New York Magazine*, (New York: Condé Nast) 23rd August.